PHYSIOLOGICAL AND BEHAVIORAL STUDIES OF SOUND LOCATION PROCESSING IN AUDITORY CORTEX OF THE MARMOSET MONKEY

by

Evan Remington

A dissertation submitted to Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy

Balitmore, MD 2013

Abstract

For many animals, the ability to localize sound sources provides a key source of information about the objects and events in the world around them. Unlike visual information, sound is not easily occluded by physical objects, and can reach the sensory epithelia from any direction, day and night. Integration of sound localization cues first occurs well below the level of auditory cortex, yet it has been well documented that auditory cortex is required for many behaviors requiring sound localization. However, the nature of spatial representation and computation in auditory cortex remains incompletely understood. This is may be due in part to the lack of studies which have measured responses from the full spatial field, and also to the lack of studies in awake, behaving animals.

In this thesis, we present data from neurons in primary auditory cortex (A1), the rostral core areas (R/RT), and the caudal belt (CM/CL) of awake marmoset monkeys (*Callithrix jacchus*) responding to broadband sounds presented from the full spatial field. The marmoset is a well established model system for the study of auditory processing and is an arboreal animal for which spatial processing is vital in its natural habitat, making it an ideal candidate for the study of spatial processing in auditory cortex. It was found that distributions of spatial receptive fields were highly heterogeneous, with neurons tuned to contralateral and ipsilateral locations, above and below the horizon, and in the rear as well as the frontal locations. Receptive field statistics varied between areas, with the caudal areas showing the most spatial selectivity, but spatial tuning was observed in neurons from all areas recorded. We also introduced a novel approach to characterize spatial tuning in a multi-source acoustic environment by presenting sounds from all sound sources simultaneously. Using this method it was found that most neurons exhibit drastically different spatial tuning in a multi-source environment compared to the single source condition, probably reflecting highly nonlinear mechanisms underlying spatial processing.

Further, we describe an auditory operant conditioning task developed in this thesis in which marmosets can be trained easily and generalize between stimulus types. This task was used to measure spatial hearing acuity (minimum audible angle) for azimuth and elevation, and found marmosets to be have acuity roughly as expected based on head size. This task is ideal for use in both head-fixed and head-free restrained neural recordings; when comparing neural responses the location discrimination task to those measured while marmosets sat passively, a subset of neurons was found with increased firing rates to one or more target locations during task engagement. Effects were similar across areas, but were largest in the caudal areas CM/CL. Together, these results suggest that spatial tuning auditory cortex of awake, behaving animals is highly selective and can be dynamically modulated to accommodate specific task demands.

Acknowledgements

I would like to extend my gratitude and appreciation to several groups of people. First, the faculty of Johns Hopkins University: I am grateful to my advisor Xiaoqin Wang, for urging me to pursue important questions and allowing me the freedom to pursue those questions that I found most interesting. I am indebted to my committee members, particularly Brad May and Eric Young for their knowledge and guidance in all things auditory, but also Ed Connor and Steve Hsiao for their valuable feedback in the later stages of my thesis preparation. I also thank my course instructors and others who helped me find my way intellectually.

Second, the many postdoctoral researchers and students that I have had the pleasure of working with over the years: Lixia Gao , Cory Miller, Michael Osmanski, Sean Slee, and Yi Zhou were all great resources, colleagues, and mentors while we shared time at Hopkins. Current and former students of Wang Lab, including (but not limited to) Dan Bendor, Lei Feng, Darik Gamble, Elias Issa, Marcus Jeschke, Luke Johnson, David Kim, and Xindong Song provided a wealth of intellectual, technical, and above all moral support.

Third, the Wang Lab, Center for Hearing and Balance, and Department of Biomedical Engineering staff: Ron Atkinson, Chris Blackledge, Jay Burns, Jenny Estes, Eric Hutchinson, Hong Lan, Shanequa Smith, and Nate Sotuyo for technical and animal care assistance.

Fourth, I thank several undergraduate assistants, in particular Jenny Kim and Mimi Gu for their help in conducting behavioral studies.

Finally, I thank my family and friends, including my parents, Mary and George, my siblings Meryl and Henry, and my wife Rebecca for all their support (and encouragement) over the years.

Table of Contents

List of Figures

Chapter 1: Introduction

Location is an intrinsic property of a sound source, and in most cases sound localization is performed unconsciously and effortlessly. Sound location information is used by the auditory system in a variety of contexts and is integral to the survival of many species: an animal may localize a sound to identify an object's location, discriminate sound locations to track objects over time, or use spatial cues to extract information from one of several sound sources. However, the nature of spatial computation and representation in the brain is a non-trivial problem. Because auditory receptor epithelia do not encode spatial information as in vision and somatosensation, spatial location in the auditory system must be computed from several acoustic cues. These computations are thought to be performed in several subcortical nuclei in the auditory pathway in a nonlinear manner (Goldberg & Brown 1969; Yu & Young 2000; Peña & Konishi 2001).

Auditory cortex, which may not be directly performing computations to determine sound location, is essential for many behaviors involving sound localization in mammals. It has been shown in many species including cats (Jenkins & Masterton 1982; Jenkins & Merzenich 1984; Malhotra & Lomber 2007; Lomber et al. 2007; Lomber & Malhotra 2008), ferrets (Kelly & Kavanagh 1987; Bizley et al. 2007), non-human primates (Heffner & Heffner 1990; Heffner 1997), and humans (Greene 1929; Wortis & Pfeffer 1948; Zatorre & Penhune 2001) that lesions and reversible deactivation of auditory cortex produce profound deficits in sound localization. It has also long been known that neurons in auditory cortex are sensitive to sound location (Eisenman 1974; Sovijärvi & Hyvärinen 1974; Benson et al. 1981; Middlebrooks & Pettigrew 1981). How the spatial field is represented in auditory cortex however, remains incompletely understood due to several limitations of past studies. Specifically, no study has previously measured responses to the full spatial field in an awake primate. Also, it is possible that the brain

employs varied processing strategies to accomplish different tasks involving spatial cues, suggesting that spatial information may be processed by the auditory system differently depending on task demands and behavioral state. Indeed, several studies have shown behavioral modulation of spatial responses in behaving animals (Benson & Hienz 1978; Benson et al. 1981; Scott et al. 2007; Lee & Middlebrooks 2011; Lee & Middlebrooks 2013). These observations are varied, and current understanding of the effects of behavior on spatial coding in auditory cortex is incomplete.

The marmoset is an attractive model system for studying auditory processing and vocal communication due to its easily accessible auditory cortex and its high vocal activity in captivity (Wang 2000). This species has been used in recent years to study coding of pitch and complex spectral features in auditory cortex (Barbour & Wang 2003b; Kadia & Wang 2003; Bendor & Wang 2005; Sadagopan & Wang 2009), temporal processing in auditory cortex (Lu, Liang & Xiaoqin Wang 2001; Bendor & Wang 2007; Kajikawa et al. 2008), thalamus (Bartlett & Wang 2007), and inferior colliculus (Nelson et al. 2009), coding at different sound intensities (Sadagopan & Wang 2008; Watkins & Barbour 2008; Watkins & Barbour 2011), auditory cortex connectivity (de la Mothe et al. 2006b; de la Mothe et al. 2006a; Reser et al. 2009; de la Mothe et al. 2012), auditory feedback mechanisms (Eliades & Wang 2008), and processing and control of conspecific communication in prefrontal cortex (Miller, Dimauro, et al. 2010). The marmoset has also recently become the first primate species in which germline expression of a transgenic modification has been achieved (Sasaki et al. 2009), broadening its potential as a model for cognitive function in disease. As a tropical arboreal species, marmosets need to navigate their environment using acoustic spatial cues. Spatial processing is therefore an important function performed by the marmoset's auditory system, making the marmoset an idea model species for further studies of spatial coding in auditory cortex.

For this thesis, we designed and built an experimental setup based on our laboratory's established single-unit recording methods for studying the representation of the full spatial field in marmosets. We also developed a simple auditory operant conditioning task which was used to measure spatial acuity in marmosets, then combined these two methodologies to study the effects of behavioral engagement on spatial coding in marmoset auditory cortex.

1.1 Characterization of full-field spatial receptive fields in an awake primate

There are two main limitations of many existing studies of spatial coding in auditory cortex. First, most studies have measured responses to sounds along single orthogonal dimensions, that is azimuth (Imig et al. 1990; Middlebrooks et al. 1994; Barone et al. 1996; Middlebrooks et al. 2002; Woods et al. 2006) and elevation (Xu et al. 1998; Stecker et al. 2003), or a limited portion of frontal space (Recanzone et al. 2000; Zhou & Wang 2012). The second limitation is the use of general anesthesia; all existing data for responses to a full- or near full-field spatial field, in addition to studies of elevation tuning, have been gathered from anesthetized animals (Brugge et al. 1996; Xu et al. 1998; Mrsic-flogel et al. 2005). These studies show a strong predominance of very broad, contralaterally and frontally biased spatial receptive fields which almost universally increase in size as sound level increases. More recent studies of azimuth (Mickey & Middlebrooks 2003; Woods et al. 2006) and partial full-field responses (Zhou & Wang 2012) have shown that responses can be restricted in space, and receptive fields on average do not increase in size as sound level increases. The effects of general anesthesia on elevation tuning are unknown. It is possible that spatial information at the level of auditory cortex may not be represented in as simple a way as suggested by these previous studies. This is underscored by patterns of non-contiguous errors in location perception (Macpherson & Middlebrooks 2000; Bremen et al. 2010). Therefore, valuable insight into spatial processing in auditory cortex could be garnered by measuring full field spatial responses in an awake animal.

Additionally, auditory cortex contains multiple well defined areas along each hierarchical processing level (Kaas & Hackett 1998; Rauschecker 1998; Hackett 2011). There is considerable evidence that these areas may have different distributions of spatial receptive fields. Several studies, both in awake and anesthetized animals and in several model species, have shown that caudal areas at multiple hierarchical levels are more selective to spatial location, on average, than primary auditory cortex or rostral areas (Rauschecker & Tian 2000; Tian et al. 2001; Stecker et al. 2003; Woods et al. 2006; Lee & Middlebrooks 2013). Behavior studies lend additional support to the hypothesis that caudal areas are important for sound localization: in cats, deactivation of a caudal auditory field, but not an anterior field, disrupted a trained azimuth localization behavior (Lomber & Malhotra 2008). Unfortunately, there are no data of full-field spatial receptive fields in the caudal or rostral areas. Thus it is not known whether differences between areas are confined to the azimuth dimension. A recent study in which responses to a partial frontal field were recorded in awake marmosets, for instance, found only subtle differences between primary auditory cortex (A1), and the caudal fields (Zhou & Wang 2012).

In natural acoustic environments, the auditory system often has to identify a sound source's location among competing sounds from multiple locations. While there is extensive data characterizing the ability of human subjects to localize and separate sounds among several spatially distributed sources, studies of spatial processing in auditory cortex have primarily relied on single source stimulus presentation. In the frequency domain, many have studied auditory cortex using complex stimuli, such as multiple pure tones (Suga et al. 1979; Kadia & Wang 2003), or sounds with random sound spectra (Klein et al. 1997; Yu & Young 2000; Slee & Young 2013). In this way, inhibition can be more easily measured in neurons with low spontaneous firing rates, and nonlinear interactions between frequency channels can be studied. The nature of representation of multiple simultaneous spatially distributed sound sources, and therefore any potential interaction between spatial locations, is largely unknown. Recent studies from our

laboratory, however, suggest that interactions between spatial locations are strong and suppressive, and with equal strength within and outside a neuron's spatial receptive field (Zhou & Wang 2012).

We measured responses of single neurons in primary auditory cortex (A1), the rostral fields (R/RT), and the caudal belt areas (CM/CL) of marmoset monkeys (*Callithrix jacchus*) to broad-band sounds presented from the free-field across the full spatial field. To do this, a sound delivery and electrophysiology apparatus was built which would allow delivery of sounds from any almost any point in the spatial field. Highly heterogeneous distributions of spatial receptive field size, shape, and location was observed in all areas tested, and receptive field shapes varied along biologically relevant dimensions. The strength of population responses was roughly flat within the contralateral and ipsilateral hemifields, with increasing contralateral bias from R/RT to A1 and from A1 to CM/CL. We also employed a systems identification approach to estimate spatial receptive fields by simultaneously presenting broad band sounds with randomized intensities from all speakers. While most neurons showed little correlation between receptive fields measured with single-speaker stimuli and multiple speaker stimuli, there was agreement for others. For these neurons, the low correlation between the two types of receptive fields suggests highly non-linear processing in integrating sounds from multiple spatial locations. These results provide the first measure of spatial responses to a full spatial field in an awake primate.

1.2 Development of an Auditory Behavior Task for the Common Marmoset and measurement of spatial acuity (minimum audible angle)

Ultimately, developing an understanding of the neural basis of perception and cognition requires the ability to link brain activity with behavior. Our laboratory has developed techniques to study natural vocal behaviors of marmosets in free moving conditions (Miller & Wang 2006; Miller et al. 2009; Miller, Mandel, et al. 2010). However, answering questions regarding the

neural basis of auditory perception often requires strict control of experimental conditions (for example, tests of spatial acuity demand a controlled head position) which is difficult to achieve in natural behavior conditions. Many animal models have well defined auditory behaviors for use in auditory physiology studies (e.g. ferret (Parsons et al. 1999; Fritz et al. 2003) macaque (Benson & Hienz 1978; Scott et al. 2007), cat (Jenkins & Masterton 1982; May & Huang 1996; Malhotra & Lomber 2007; Lee & Middlebrooks 2011), and rat (Otazu et al. 2009)), as do many other species for behavioral studies (e.g. horses (Heffner & Heffner 1984), gerbils (Heffner & Heffner 1988b), pigs (Heffner & Heffner 1989), cows and goats (Rickye S Heffner & Heffner 1992)).

Previously, a conditioned avoidance task was used to measure absolute hearing thresholds in marmosets (Seiden 1957). There have also been a number of studies using operant conditioning behaviors to study visual cognition in marmosets (Miles & Meyer 1956; Roberts et al. 1988). We have developed an auditory operant conditioning task for the common marmoset. Subjects must lick at a feeding tube (equipped with an infrared photo-beam) during target sound presentation in order to receive a food reward, while withholding licking when a target sound is not being presented. Most animals learned this behavior quickly and behaved consistently for relatively long periods of time. The task has already been employed in the measurement of a marmoset audiogram (Osmanski & Wang 2011) and harmonic resolvability and pitch perception (Osmanski et al. 2013). Here a complete description of the task and training procedures, additional considerations for marmoset training and behavior performance, and learning curves for 5 marmosets trained on this task are presented. Crucially, it is also shown that this task can be coupled with single-unit electrophysiology recording without causing significant interference to the recording stability. Data is presented from an animal performing a sound location discrimination task while the single-unit recordings were conducted. This work was published previously (Remington et al. 2012).

As in all mammals, sound location perception is determined by three cues: interaural time difference (ITD), interaural level difference (ILD), and spectral shape. These cues are the result of the geometry of the head and ears: the distance between the ear canals determines ITD, the size and shape of the head (and to an extent the neck and shoulders) determines ILD, and the shape of the pinna (or outer ear) modulates the shape of the incoming sound spectrum by introducing resonances and notches in a spatially dependent manner. This spatially dependent acoustic filter is referred to as the head related transfer function (HRTF). This creates a dichotomy of perceptual computations for spatial perception: ITD and ILD are binaural cues and provide, at least in mammals, useful information about an object's lateral position. A meta-analysis of measurements in many animal species found that at least for non-echolocating animals, there is a roughly linear relationship between head size and horizontal acuity (Brown & May 2005), although there is also evidence that animals with highly focused binocular vision may also be localization specialists (R S Heffner & Heffner 1992). Spectral cues, conversely, provide information related to front/back and up/down localization. It is appropriate therefore to measure acuity along these two axes independently. In this study, we measured the minimum audible angle (MAA) of broad band sounds in azimuth and elevation. The results indicate that marmosets' horizontal and vertical spatial acuity is roughly on par with other species of similar size previously tested.

1.3 Representation of space in an actively behaving primate

It is well known that response properties of neurons in auditory cortex can change rapidly to adaptively match specific task demands. These phenomena have been studied extensively for frequency tuning, in which neurons' receptive fields change in ways which are highly tailored to relevant task stimuli (Fritz et al. 2003; Fritz et al. 2007). Similar effects have been found in tasks requiring attention to temporal sound attributes (Fritz, Elhilali, et al. 2005; Niwa et al. 2012).

Auditory cortex is essential for many behaviors involving sound localization in mammals (Jenkins & Masterton 1982; Bizley et al. 2007; Lomber & Malhotra 2008), so it is of great interest how spatial processing in auditory cortex is affected by behavioral context. However, representation of the spatial field in auditory cortex at the level of single neurons during active listening remains largely unclear. The few studies which have compared spatial coding in different behavioral contexts have shown varied effects of task engagement on responses in auditory cortex; however only some of these changes appeared to be beneficial neural adaptations. In one study where macaques were asked to respond only to sounds presented to an attended ear, responses of a minority of contralateral ear-preferring neurons increased to the target location when the target location was the contralateral ear (Benson & Hienz 1978). However, the same group later found that while subjects performed a free-field localization task, responses increased to selected locations arbitrarily in a subset of neurons (Benson et al. 1981). More recently, responses were shown to be increased while subjects performed an interaural phase discrimination task, although these changes did not confer an overall increase in neurometric discrimination thresholds (Scott et al. 2007). Finally, a study in cats showed that responses to background sounds at non-preferred locations decreased, resulting in narrower tuning widths while subjects performed a location discrimination task (Lee & Middlebrooks 2011; Lee & Middlebrooks 2013).

Based on the wide range of effects seen in previous studies, it seems clear that a more complete understanding of spatial coding in auditory cortex subserving behavior requires additional studies measuring activity in auditory cortex during active listening conditions. One limitation of previous studies is that their results are interpreted in terms of receptive fields constructed by sampling various limited regions of space. As mentioned previously, little is known regarding full-field spatial receptive fields in awake, passive subjects, let alone during behavioral engagement. It is possible that the effects of active behavior can be better

contextualized by measuring full-field spatial receptive fields in the passive condition as well as responses to a large spatial area in behavior. Also, as mentioned previously, many studies have shown that the distributions of spatial tuning properties vary between auditory areas along the rostral-caudal axis, with neurons in caudal areas, on average, displaying higher selectivity for spatial locations than those in rostral areas and primary auditory cortex (Rauschecker & Tian 2000; Stecker et al. 2003; Woods et al. 2006; Lee & Middlebrooks 2013). However, only one study (Lee & Middlebrooks 2013) has compared the effects of engagement in a spatial auditory task between primary, rostral, and caudal areas. We therefore asked whether any behavior effects would differ qualitatively or quantitatively along the rostral-caudal axis.

Single-unit responses were recorded in auditory cortex of marmosets while subjects performed a spatial discrimination task in different regions of the spatial field. Comparing these responses to those measured while marmosets sat passively, it was observed that a subset of neurons increased firing rates to one or more target locations during task engagement. The effect of these firing increases served to increase firing rates relative to the reference location. Increases occurred both within and outside of the classically defined spatial receptive field, but were moderately larger when responses at the target location were greater than responses at the reference location. Comparing effects of behavior between rostral (R/RT), caudal (CM/CL), and primary (A1) auditory areas, the largest effects were found in CM/CL and the smallest in R/RT. These results add support to the hypothesis that sound-source locations are represented by cortical populations whose responses can be modulated to best suit the demands of a particular task (Lee & Middlebrooks 2011).

Chapter 2: Materials and Methods

2.1 Animal preparation and electrophysiological procedures

A chronic recording preparation (Lu, Liang & X Wang 2001) was used to record single-neuron activity in the auditory cortex (left hemisphere) of three awake common marmoset monkeys (Callithrix jacchus). All subjects were trained to sit in a custom-designed primate chair and perform a behavioral task (Remington et al. 2012) for a separate study. After 1-2 months of behavioral training, two stainless steel headposts were attached to the skull under sterile conditions with the animal deeply anesthetized by isoflurane (0.5–2.0%, mixed with 50% O2 and 50% nitrous oxide). The headposts served to maintain a stable head orientation of the subject during electrophysiological recordings, however in this study, only one (the front) headpost was used for head fixation. To access auditory cortex, small craniotomies (1.1 mm in diameter) were made in the skull over the superior temporal gyrus to allow for the penetration of electrodes (tungsten electrodes, 2- to 5-MΩ impedance, A-M Systems, Carlsborg, WA) via a hydraulic microdrive (Trent-Wells, Los Angeles, CA). Single unit spiking activity was sorted online using a template-based spike-sorting program (MSD, Alpha Omega Engineering) and analyzed using custom programs written in Matlab (Mathworks, Natick, MA). Experimental procedures were approved by the Institutional Animal Care and Use Committee of the Johns Hopkins University following National Institutes of Health guidelines.

2.2 Experimental setup

Experiments were conducted in a double-walled sound-attenuated chamber (Industrial Acoustics, IAC, New York) with the internal walls, ceiling, and floor were lined with ~3 inch acoustic absorption foam (Sonex, Illbruck). Acoustic stimuli were delivered using an array of 24 speakers

(FT28D, Dome Tweeter, Fostex) covering a complete sphere (Figure 3.1). The loudspeakers were mounted at a distance of 1 m to the animals head and covered 5 Elevations (ELs) at 45° spacing and several Azimuths (AZs). One speaker was located directly above the animal, 7 speakers each were evenly spaced at ± 45° EL (AZ at –45° EL: ±25.7°, ±77.1°, ±128.6° and 180°; AZ at 45° EL: 0°, ±51.4°, ±102.9° and ±154.3°), 8 speakers were evenly positioned at 0° EL (AZ: 0°, ±45°, ±90°, ±135°, 180°), and finally 1 speaker was located at either −85° EL (for subject M3T) or −67.5° EL at 0° AZ (for M71V and M9X). Subjects sat in a wire mesh primate chair mounted onto a single stainless steel bar such that the animals head was centered in the room. In this text positive azimuth angles correspond to speakers ipsilateral to the recording site or to an ipsilateral shift if changes in azimuth were analyzed. During experiments eye position was not controlled.

2.3 Acoustic stimuli

Stimuli were generated in Matlab at a sampling rate of 97.7 kHz using custom software. For single channel delivery, digital signals were converted to analog (RX6, 2-channels D/A, Tucker-Davis Technologies, Alachua, FL), then analog signals were attenuated (PA5 x2, Tucker-Davis Technologies), power amplified (Crown Audio x2, Elkhart, IN), and played through a chosen channel of a power multiplexer (PM2R x2, 16-channels, Tucker-Davis Technologies). For multi-channel sound delivery, digital signals were attenuated, converted to analog (RX8, 23-channels D/A, and RX6, 1-channel D/A, Tucker-Davis Technologies), and power amplified (PA8 x3, 8 channel power amplifier, Tucker-Davis Technologies). Loudspeakers had a relatively flat frequency response curve (± 3-7 dB) and minimal spectral variation across speakers (< 7 dB re mean) across the range of frequencies of the stimuli used; all large (5-7 dB) spectral deviations occurred in narrow bandwidths near the upper limit of speakers' frequency range (above 28 kHz), above the first spectral notch measured in marmoset head related transfer functions (Slee & Young 2010). Stimulus intensity was calibrated to 95 dB SPL at 0 dB attenuation for a 4 kHz tone. All intensities in this report are expressed in terms of the peak-to-peak pure-tone equivalent

dB SPL. When possible, neurons were characterized for frequency, intensity, and spatial tuning. For frequency tuning, stimuli consisted of pure tones, band-pass filtered Gaussian noise, Random Spectral Shape (RSS) stimuli (Yu & Young 2000; Barbour & Wang 2003a), and occasionally frequency modulated (FM) sweeps. The frequency axis was sampled in 0.1 octave steps, typically over a 4 octave range (2-32 kHz). All firing rates were calculated over a time window beginning 15 ms after stimulus onset and 20ms after stimulus offset; "offset" responses (occurring > 20ms post stimulus) were not analyzed. Best frequency was defined as the frequency which led to the maximum evoked significant firing rate, or for neurons only driven by RSS stimuli, the highest calculated linear RSS weight (in spikes/dB/second). For spatial tuning, stimuli included band pass filtered unfrozen Gaussian noise, single RSS stimulus tokens, and occasionally FM sweeps. All stimuli used to measure spatial receptive fields (SRFs) were either band-pass filtered or constructed to have energy between 2 and 32 kHz. Stimuli were typically 200ms long with 10ms cosine ramps and delivered in pseudorandom order, and, with the exception of RSS and RLS stimulus sets, delivered between 5 and 10 times.

2.4 Characterization of spatial receptive fields with a single sound source

Spatial receptive fields were generated by interpolating responses to the 24 speaker array into a 5˚x 5˚ (2592 zone) vertical pole grid according to the spherical distance-weighted mean of the two nearest speaker locations, with the weight at a single interpolated location from speaker *i* calculated as:

$$weight_i = \frac{(1/distance_i)^2}{\sum_{j=1}^{2}(1/distance_j)^2} \tag{2.1}$$

Where distance is the spherical distance between speaker *i* and the center of the grid zone. This interpolated response map was then used to calculate several tuning properties. First,

$$tuning\ area = \frac{\sum_{\{i:r_i>thresdhold\}} area_i}{total\ area} \tag{2.2}$$

where $area_i$ is the area of grid zone i, r_i is the interpolated firing rate for grid zone i, and *threshold* is classically defined as the half maximal firing rate. The tuning vector (TV) roughly characterizes neurons' preferred direction as well as receptive field dispersion (tuning vector length). It is calculated as following:

$$tuning\ vector = \frac{\sum r_i * vector_i * area_i}{\sum r_i * area_i} \tag{2.3}$$

Where $vector_i$ is the unit vector pointing to the center of grid zone i. A response to a single location results in tuning area of ~.045 and tuning vector length of ~1.

Lateral and median centroids were calculated in a similar manner to the tuning vector, with responses averaged across the orthogonal dimension. Negative values indicate contralateral and lower hemifields for lateral and midline angles, respectively. For lateral and median centroid calculations, firing rates were re-interpolated into a horizontal pole coordinate system. Lateral centroid angles range from -90 to 90°, while median plane centroid angles ranged from 0 to 360°. SRFs are plotted using a vertical pole coordinate system. Finally, the "number of peaks" was the number of contiguous regions above threshold.

2.5 Characterization of spatial receptive fields with random spatial profile (RSP) stimuli

Spatial receptive fields in a spatially dense acoustic environment were characterized by playing sets of broad band sounds from the entire 24-speaker array simultaneously, randomizing the sound level from each speaker. This stimulus delivery paradigm was adapted from a similar method (with stimuli referred to as "random spectral shape" stimuli) used to study spectral processing in the auditory system (Yu & Young 2000; Barbour & Wang 2003a; Slee & Young 2013). The complete set of stimuli comprised a stimulus matrix $\mathbf{\Lambda}$ of mean-subtracted intensities in which the rows represent individual stimuli and the columns represent the individual speakers.

This RSP stimulus set was constructed to sample the space of all possible 24-location spatial profiles in such a way that responses to the stimuli could be used to construct linear model of the spatial tuning to spatially dense stimuli. To do this, the set was constructed such that the levels of each speaker were statistically independent across all stimuli. This condition is satisfied if the location intensity autocorrelation matrix is equal to the identity matrix:

$$\mathbf{\Lambda}^{\mathrm{T}}\mathbf{\Lambda} = n\sigma^2_{\Lambda^{\mathrm{T}}\Lambda}\mathbf{I} \tag{2.4}$$

where n is the number of stimuli and $\sigma^2_{\Lambda^{\mathrm{T}}\Lambda}$ is the variance of the mean-subtracted stimulus levels. This constraint can only be met for a matrix having more rows than columns. Therefore, the minimum number of stimuli required to construct a linear estimate of the spatial weighting function in this case was 25. In previous studies, at least one flat stimulus is also included, giving total of 26 stimuli. The linear model can be written as:

$$\vec{R}_{lin} = R_0 + \mathbf{\Lambda}\vec{w} \tag{2.5}$$

where $\vec{R}_{lin}$ is a column vector of m rate values predicted in response to a set of m different RSP stimuli, R_0 is the firing rate to an RSP stimulus with a flat spatial profile, and $\vec{w}$ is the 24 value linear weighting vector. This equation is referred to as the linear synthesis equation. The weighting function is calculated as:

$$\vec{w} = \left(\mathbf{\Lambda}^{\mathrm{T}}\mathbf{\Lambda}\right)^{-1}\mathbf{\Lambda}^{\mathrm{T}}\left(\vec{R} - R_0\right) = \frac{\mathbf{\Lambda}^{\mathrm{T}}\vec{R}}{n\sigma^2_{\Lambda^{\mathrm{T}}\Lambda}} \tag{2.6}$$

where $\vec{R}$ is the measured firing rate vector (in spikes/second) to the RSP stimulus set. Weights $\vec{w}$ are expressed in units of spikes/second/dB. This is equation is referred to as the analysis equation. Thus, when the intensity is increased at a location with a positive weight function, firing rate should increase as well. A more complete description of these methods can be found in previous work (Barbour 2002).

While the required number of stimuli to independently estimate linear weights is 26, estimations improve when the size of the stimulus set increases (Yu 2003). Neurons were characterized with sets of 26, 51, or 255 RSP stimuli with 5-10 repetitions for sets of 26 and 51 stimuli, and 1 or 2 repetitions with sets of 255 RSP stimuli. The intensity standard deviation parameter σ was 5, 10, or 15dB, and mean level was 40, 50 or 60dB attenuation. Units characterized with sets of 26 or 51 stimuli were determined to be significantly driven if at least one stimulus resulted in a statistically significant driven rate, as measured with a Wilcoxon rank-sum test at $\alpha = 0.001$. Due to the difficulty in achieving significance to individual stimuli when low stimulus repetition numbers are used, units characterized with sets of 255 stimuli were determined to be significantly driven if the null hypothesis that the firing rate distribution during the stimulus period was identical to the firing rate distribution during the pre-stimulus period was rejected using a Kolmogorov-Smirnov test with $\alpha = 0.001$. Units with no significant response to RSP stimuli were not analyzed further.

2.6 Comparison of single-speaker and multiple speaker receptive fields

Because single speaker spatial receptive fields are represented with raw firing rates while RSP receptive fields are represented by weighting functions (in units of spikes/second/dB), a direct comparison between firing rates cannot be made. Instead we calculated Pearson's correlation coefficient between the two vectors to quantify how well the two receptive fields agreed.

RSP linear model fit. If the linear model describes neurons' behavior, it should be able to predict responses to stimuli that are not part of the ensemble used to compute the weights. Weighting functions were tested for goodness of fit by calculating a quality factor which was based on the mean squared error between the predicted and observed rates divided by the variance of the predicted rates (Yu & Young 2000). To avoid overfitting, individual rates were predicted using a

weighting function calculated with that stimulus/response pair removed from the set. The quality factor was calculated using the following equation:

$$Q = \frac{\sum_i \left(\hat{r}_i - \bar{\hat{r}}\right)^2}{\sum_i \left(r_i - \hat{r}_i\right)^2 + \sum_i \left(\hat{r}_i - \bar{\hat{r}}\right)^2} = \frac{1}{1 + \sum_i \left(r_i - \hat{r}_i\right)^2 \Big/ \sum_i \left(\hat{r}_i - \bar{\hat{r}}\right)^2} \tag{2.7}$$

2.7 Principal component analysis, clustering and Euclidean distance

In order to quantify the variability in SRF shapes, principal component analysis (PCA) was performed to identify the orthogonal dimensions along which SRFs varied most strongly. The principal component scores for the first 5 principal component dimensions, as well as the full 24-dimensional data were then used to attempt to cluster the data using a k-means clustering algorithm available in Matlab (Mathworks, Natick, MA). This is an unsupervised clustering algorithm, and will cluster data regardless of the existence of true clusters. To test for evidence for the existence of clusters within the data, the gap statistic (Tibshirani et al. 2001) was measured on the results of the k-means clustering algorithm for each number of clusters. First, let

$$W_k = \sum_{r=1}^{k} \frac{1}{2n_r} D_r \tag{2.8}$$

with D_r representing the sum of all pairwise Euclidean distances in cluster r. W_k represents the within cluster sum of squares summed over all clusters. The gap statistic is then calculated as

$$Gap_n(k) = E_n\{\log(W_\mathrm{k})\} - \log\left(W_k\right) \tag{2.9}$$

where n is the number of samples in the population, k is the number of clusters imposed on the dataset, and $E\{\log(W_\mathrm{k})\}$ is the expectation of W_k with a sample of size n using an appropriate reference distribution. The goal is to maximize the "gap" between the reference W_k and the actual W_k by correctly choosing k, the number of clusters. The reference distribution was a scaled

multivariate uniform distribution. For a complete description of the Gap statistic calculation, see (Tibshirani et al. 2001).

The dissimilarity of the population responses to the 24 speaker locations by was quantified by calculating the Euclidean distance of the n-neuron response vectors for each pair of locations, normalized by the number of neurons in the population.

$$\vec{d}_{i,j} = \sqrt{\frac{1}{n}\sum_{n}\left(r_i^n - r_j^n\right)^2} \tag{2.10}$$

where r_i^n and r_j^n are the averaged responses of neuron n to locations i and j for a constant stimulus type and intensity. For each speaker location, the Euclidean distances to the 4 closest locations (in terms of population response) were averaged as a measure of response uniqueness for that location.

$$d_i = \frac{1}{4}\sum_{j=\min\,(d_{i,j})}^{4} d_{i,j} \tag{2.11}$$

2.8 Identification of A1, the rostral fields R/RT, and caudal areas CM/CL

In this study, we recorded responses from several areas, including A1 (primary auditory cortex), the rostral areas R (for rostral) and RT (for rostrotemporal), and the caudal belt areas CL (caudolateral) and CM (caudomedial). In the marmoset, similar to other nonhuman primate species, A1 is situated largely ventral to the lateral sulcus on the superior temporal plane and exhibits a low-to-high topographical frequency gradient along the rostral-caudal axis. The boundary between A1 and the rostral field R can be identified by a downward to upward frequency gradient reversal along the rostral-caudal axis. Conversely, areas CL and CM can be identified by an abrupt decrease of best frequency at the high frequency (caudal) border of A1 (Merzenich & Brugge 1973; Aitkin et al. 1986; Kaas & Hackett 2000). In this study, the boundaries between R/RT and A1 and A1 and CM/CL were set by plotting the average best

frequency along the rostral-caudal axis, approximately parallel to lateral sulcus, and setting a hard boundary at the local minimum and maximum, respectively, between the two areas. Frequency maps and corresponding reversals separating recorded areas are shown in figure 2.1. In the two animals for which recordings were made in CM/CL, recordings proceeded rostrally until units were no longer driven by auditory stimuli, and were sometimes driven by moving visual targets (tested using a laser pointer). Neurons were not separated further into R/RT or CM/CL in this study.

2.9 Behavior apparatus

The operant behavior setup included the restraining chair, a behavior response apparatus, a reward delivery system, and a stimulus delivery and behavior control system. Marmosets made behavior responses by licking at a feeding tube; responses were measured by a custom built lick detector which registers whether an infrared beam in front of the animal's mouth has been interrupted. When the animal's head is not restrained, this can also be accomplished by moving its face into the detector. A programmable syringe pump (NE-500, New Era Pump Systems, Wantagh, NY) delivered food reward through a disposable IV extension and into a custom machined lexan tube which was positioned via a custom machined bracket fastened to the neck plate. Reward was a mixture of single-grain rice cereal (Gerber), strawberry and/or banana-flavor (Nesquik), a protein powder supplement (Nutiva), and baby formula (Similac). This mixture is nutritionally substantial and of relatively low viscosity for pumping efficiency; a single reward was between 0.1 and 0.2 ml and could be delivered within a few seconds. The speed of delivery is limited by mixture viscosity and pump speed and power.

A computer running custom software written in Matlab and the RX6 multifunction processor controlled behavior: Matlab software controlled stimulus generation and behavior flow, while the RX6 unit served to synchronize stimulus delivery, reward delivery, behavior responses and

single-unit electrophysiology data (when applicable). A custom built power/relay module powered and electrically isolated the computer from the equipment inside the experimental chamber. The marmoset chair and feeding tube, along with a system schematic, are illustrated in Figure 4.1.

All behavior animals were maintained at approximately 90% of their free-feeding weight on a diet consisting of a combination of monkey chow, fruit and yogurt and had ad libitum access to water. For measurement of perceptual thresholds, subjects were tested once a day, five days per week between the hours of 0900 and 1800. During training and testing, animals were monitored by closed circuit infrared camera.

2.10 Measurement of perceptual thresholds for location discrimination

To quantify behavior performance for the measurement of minimum audible angle in marmosets, corrected hit rate was calculated:

$$HRc = \frac{HR - FAR}{1 - FAR} \tag{2.12}$$

where HR is the raw hit rate, and FAR is the false alarm (or false positive) rate. Using this measure, psychometric functions were generated and thresholds were determined to be point at which the interpolated function equaled 50%. Final thresholds were calculated from psychometric functions averaged over 3 of 4 consecutive sessions during which thresholds were no higher than the mean plus 5° and during which time thresholds did not appear to be systematically decreasing.

2.11 Effects of behavior on spatial responses

Four target/background configurations were used. The background location was 45° off of midline (front and back, contralateral and ipsilateral), and the target locations were the most lateral positions (±90°, same in all conditions), and also 45° above and below the horizon, but in

the same azimuthal quadrant as the background location. For targets above and below the background location, their azimuth locations were either 51° or 25.5° off of the midline (one of each per condition). A diagram of one such condition is shown in figure 5.1. Sound level was roved either ±5 dB SPL or ±10dB SPL depending on each neurons' sensitivity to sound level.

Three measures were calculated to quantify the difference in firing rate between behaving and passive conditions. Modulation index (MI) is a measure of the response difference at a particular location scaled by the combined response strength at that location:

$$MI = \frac{R_B - R_P}{R_B + R_P} \tag{2.13}$$

where R_B is the firing rate in the behaving condition and R_P is the firing rate in the passive condition. Second, the scaled shift is a measure of the response shift at a location relative to maximum passive firing rate:

$$S_R = \frac{R_B - R_P}{max(R_P)} \tag{2.14}$$

where $max(R_P)$ is the maximum firing rate among all locations in the passive spatial receptive field. Finally, the shift in the difference between the firing rates to each target speaker and the background location and the difference between those same locations in the passive condition was calculated:

$$\frac{(R_B^T - R_B^B) - (R_P^T - R_P^B)}{max(R_P)} \tag{2.15}$$

where R^T is the firing rate to a target location, and R^B is the firing rate to the background location.

Rates during active behavior were compared to two different behavior conditions. The first condition is the standard spatial receptive field measured by playing sounds from all 24 speaker locations in a randomized order. To control for effects of stimulus order, which can result

in suppression or facilitation of neural responses (Ulanovsky et al. 2004; Bartlett & Wang 2005), firing rates were compared in the behavior condition to a second passive condition in which stimulus delivery was identical to the behavior condition except that stimuli did not stop if behavioral responses were measured. If subjects, which were being monitored via closed circuit television, responded during this control condition, the data were rejected. To reduce confusion, subjects were cued to the beginning of behavior sessions by alternating the house light on and off.

2.12 Statistical tests

In the following studies, all statistical comparisons between pairs of data sets are made using a Wilcoxon rank-sum test, and all stated correlations are Pearson's correlation coefficient (Pearson's r) with a corresponding p values representing student's t-tests on a transformation of the correlation.

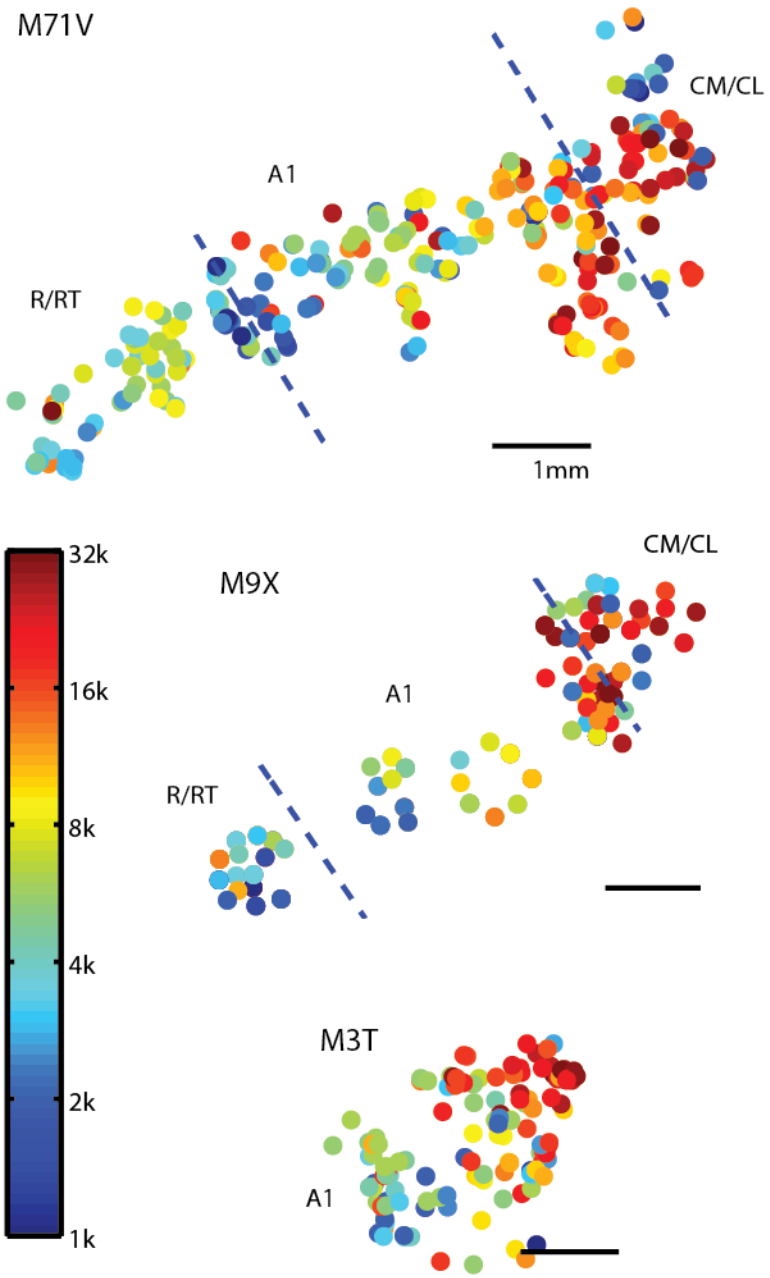


Figure 2.1: Best frequency maps and identification of auditory cortical areas. The boundary between A1 and the rostral field R can be identified by a downward to upward frequency gradient reversal along the rostral-caudal axis; areas CL and CM can be identified by decrease in best frequency at the high frequency (caudal) border of A1 (Merzenich & Brugge 1973; Aitkin et al. 1986; Kaas & Hackett 2000).

Chapter 3: Full-field spatial receptive fields in auditory cortex

3.1 Results

3.1.1 Distribution of spatial receptive fields in auditory cortex.

We recorded well-isolated single unit activity in the auditory cortex of 3 marmoset monkeys while broad-band sounds were played from a 24 speaker free-field array (figure 3.1). During recording sessions, marmosets sat passively in a wire mesh chair which was designed to minimize acoustic reflections near the pinna. In total, spiking activity was recorded from 648 units (408 in A1, 143 in CM/CL, and 97 in R/RT) which responded with a significantly elevated firing rate above the spontaneous rate to one or more locations ($p < 0.001$, Wilcoxon rank-sum test, minimum 1 spike/stimulus). Figure 3.2 shows an example neuron which responded to sounds only at the most contralateral locations. At its preferred location this neuron responded with a sharp burst of spikes and a weaker sustained response; away from its preferred location it fired weakly or not at all. Neurons in auditory cortex displayed a wide variety of spatial response patterns. Spatial receptive fields were centered in contralateral, ipsilateral, lower, upper, front, and rear regions of space, with some neurons displaying multi-peaked receptive fields. Figure 3.3.1-22 shows several examples of different types of spatial receptive fields.

Distributions of best locations and tuning vectors in figure 3.4.1 show that in marmoset auditory cortex, receptive fields covered all of auditory space. The majority of neurons responded most strongly to locations in the contralateral hemifield, a trend that was strongest in CM/CL and weakest in R/RT. Within hemifields, however, representation of space was broadly distributed across neurons, with relatively little variation in the population firing rate strength (figure 3.4.2). Tuning vectors tended to be centered at more contralateral locations, with 148/408**,** 58/143, and

28/97 neurons in A1, CM/CL, and R/RT respectively with scaled lateral tuning vector component <−0.8 (with −1 signifying a tuning vector pointing to 90° contralateral). However, this bias became very weak when the correlation between laterality and individual firing rates in the contralateral hemifield (including the median plane) was considered (A1: $r = 0.04$, $p < 0.002$; CM/CL: $r = -0.02$, $p = 0.3$; R/RT: $r = 0.008$, $p = 0.8$). This discrepancy can be explained by a portion of neurons with vector averages centered at very contralateral locations while not firing strongly at the most contralateral locations (for example, a neuron which fires to ~45 degrees contralateral locations, but has little median plane sensitivity). Additionally, there was a significant number of ipsilateral preferring neurons, and these too were distributed throughout the hemifield. Overall, we find that auditory space, particularly in the contralateral hemifield, is broadly represented.

3.1.2 Distribution of tuning selectivity in auditory cortex

We also found that a large number of neurons, particularly in A1 and CM/CL, had spatial receptive fields covering less than 25% of space (A1: 206/408; CM/CL: 103/143; R/RT: 35/97, figure 3.5). There was a significant effect of auditory area on tuning area (ANOVA: $p << 0.001$), with R/RT having the largest tuning areas (mean = 0.43), followed by A1 (mean = 0.34), and CM/CL (mean = 0.22). As the distributions of tuning areas found here are lower than those found in some previous studies of azimuth tuning in awake animals (Lee & Middlebrooks 2011; Mickey & Middlebrooks 2003), we asked whether the measurement of tuning area could be affected by restricting data analysis to the azimuth dimension at 0° elevation. We found that in all areas, tuning areas were significantly larger when only azimuth was considered compared with those when the full spatial field was considered (mean, full /azimuth only, A1: .34/.45, $p << 0.001$; CM/CL, .22/.31, $p << 0.001$; R/RT: .43/.58, $p << 0.001$). Additionally, when only azimuth locations were considered, a significant fraction of neurons failed to reach criterion to be considered driven by broad band sounds (A1: 86/408; CM/CL: 30/147; R/RT: 18/97). Another

measure of directional selectivity is the tuning vector magnitude, with a large tuning vector signifying a highly directional neuron. Tuning vector magnitude is sensitive to shape and dispersion as well as extent of spatial receptive fields. For example, a neuron firing equally strongly to two locations opposite each other in space, and not at all to other locations, would have a small tuning area (~0.1), but also a tuning vector magnitude of 0. Thus, tuning area and tuning vector magnitude represent a complementary set of measurements to characterize spatial selectivity. The tendency of neurons to have relatively small tuning vector magnitudes suggests that many neurons are not tuned sharply in both azimuth and elevation. Though narrow tuning area does not guarantee a large tuning vector magnitude, we did find that tuning vector magnitude followed the same trend vs. area as tuning area (ANOVA: $p << 0.001$), with CM/CL having the most directional neurons (mean = 0.35), followed by A1 (mean = 0.27), and R/RT (0.16).

3.1.3 Spatial receptive fields are largely invariant to changes in sound level

Studies of spatial tuning in anesthetized animals have found that spatial receptive fields, almost universally, drastically increase in size with increasing sound level (Brugge et al. 1996; Stecker et al. 2003; Mrsic-flogel et al. 2005). Studies in of azimuth and partial field tuning in awake animals have found this effect to be weaker or nonexistent (Eisenman 1974; Mickey & Middlebrooks 2003; Woods et al. 2006; Zhou & Wang 2012). Still, it is possible that receptive fields may increase in size in an awake animal when the full spatial field is considered. In this study, there was a weak but significant effect of sound level on population tuning area in areas A1 ($r = 0.15$, $p < 0.001$) and R/RT ($r = 0.13$, $p = .05$), but not in CM/CL ($r = -0.022$, $p = 0.7$; figure 3.7**)**. Effects of sound level on the complementary measure of tuning vector magnitude were similar: there was a weak but significant effect of sound level on tuning vector magnitude in area A1 ($r = -0.1$, $p < 0.01$; figure 3.7), but there was no effect in CM/CL or R/RT. In single neurons, we measured the correlation between tuning area and sound level referenced to best SRF level (determined by highest single firing rate across all SRFs). When considering within neuron

effects, correlations were not found between ΔdB SPL and tuning area shift in areas A1 ($r = 0.10$, $p = .08$) or R/T ($r = -0.07$, $p = .4$), but there was a significant negative correlation between tuning area and sound level in CM/CL units ($r = -0.41$, $p < .001$; figure 3.8). Data points at ΔdB SPL = 0 represent instances when the same stimulus level was tested multiple times in the same neuron. No correlation was seen in any area between ΔdB SPL and tuning vector magnitude (figure 3.8). To quantify tuning area stability across level, we tested for an effect of the absolute value of level changes on the absolute value of tuning area change, both referenced to the best SRF. Here, there was no significant effect in A1 ($r = 0.015$, $p = 0.8$) or R/RT ($r = 0.05$, $p = .6$) and a weak effect in CM/CL ($r = 0.26$, $p = 0.01$) as expected due to the negative correlation between tuning area change and ΔdB SPL (figure 3.9). For tuning vector magnitude, there was a weak effect of ΔdB SPL magnitude in A1 only ($r = 0.15$, $p < .01$, figure 3.9). In general, changes in tuning area and vector magnitude were small, typically 0.1 or less, across all areas and ΔdB SPL values, and likely the result of stimulus variability rather than stimulus parameter changes.

We further tested stability of spatial tuning directionality as a function of sound level, measuring the change in tuning vector (in degrees) as a function of ΔdB SPL (figure 3.10). Here there were weak effects of ΔdB SPL on tuning vector shift magnitude in A1 ($r = 0.17$, $p < 0.01$) and R/RT ($r = 0.27$, $p < 0.01$), but not CM/CL ($r = 0.0072$, $p = 0.5$), indicating that like tuning area changes, tuning vector changes were more strongly governed by response variability. An analysis of covariance showed an effect of area on tuning vector stability, with CM/CL having the smallest tuning vector shift magnitudes (ANCOVA, $p < .05$). Figure 3.11.1-3 shows several examples of spatial receptive fields measured in single neurons across multiple sound levels.

3.1.4 Decomposition of spatial receptive fields by biologically relevant components

The distribution of tuning vectors (figure 3.4.1) suggests a broad distribution of spatial preferences, but the tuning vector incompletely describes SRF shape. Therefore, we performed a

principal component analysis to characterize the dimensions along which spatial receptive fields varied most strongly. For this analysis, data from animals in which all three areas were recorded were included (2 animals; A1: 245 units; CM/CL: 147 Units; R/RT: 97 units). The first five resulting principal components are plotted in 3.12. The first principal component, accounting for 40% of total variance, approximates the non-directional component of neural responses, and is referred to as the spatial response contrast. The dominance of this component mirrors the broad distribution of tuning area (figure 3.5). The second and third components, reflecting a neurons' hemifield and laterality preferences, together describe azimuth tuning. The fourth and fifth components represent front/back and up/down variation. Variation in azimuth makes up a larger fraction of variance than median plane tuning variation (18 % vs. 10%), while the remaining components (6-24) were less biologically intuitive and constituted the final 32% of variance.

Previous studies have found that binaural sensitivity varies continuously throughout auditory cortex (Campbell et al. 2006). We applied k-means clustering on both the first 5 principal components as well as the full 24-dimensional data set and used the gap statistic method (Tibshirani et al. 2001) to analyze the results. Evidence for clusters was not found (largest gap change: 1 cluster). Though there were no clear clusters in SRF shape across the entire population, it was still a possibility that spatial receptive fields could be quantitatively different across areas. We trained a series of linear support vector machine classifiers in "one vs. rest" configurations to classify populations from each area against the remaining two. Classification performance was generally poor using full spatial receptive fields (accuracy: 0.50 for A1; 0.64 for CM/CL; 0.61 for R/RT) and principle components 1-5 (accuracy: 0.50 for A1; 0.66 for CM/CL; 0.58 for R/RT). These results suggest that while spatial receptive fields vary along biologically relevant dimensions, spatial selectivity varies in a continuous manner, and there is significant overlap in spatial receptive field shape between areas in auditory cortex.

3.1.5 Euclidean distance between population responses to sound location

Next, we quantified the theoretical ability of neural populations in each auditory area to accurately identify sound source locations. We hypothesized that if CM/CL is important for sound localization, population firing rates to different locations should be more separable, leading to better neural discrimination. Euclidean distance was calculated, normalized by neural population size, of the mean population firing rate between each pair of locations for each auditory area. Again, only analyzed data from animals with all three areas recorded was analyzed. Figure 3.13 shows the resulting Euclidean distances, averaged across the 4 "nearest" locations (as measured by population response space rather than physical space) for each location, for each area. Moving from R/RT to A1, to CM/CL, it was found that Euclidean distance between the nearest locations increased in the contralateral hemifield and decreased in the ipsilateral hemifield (repeated measures ANOVA, $p << .001$). In the median plane, there was a significant effect of area (repeated measures ANOVA, $p << .001$), but there was no significant difference between A1 and CM/CL (repeated measures ANOVA, $p = .97$). Thus, while population responses in CM/CL are better separated for sound source locations in the contralateral hemifield; they do so at the expense of population response separation for ipsilateral locations.

3.1.6 Relationship between spatial and non-spatial tuning properties

Sound localization requires integration of several sound source attributes (i.e. binaural and spectral cues), so it is possible that spatial tuning could be correlated with non-spatial tuning properties. In marmosets, interaural level differences increase with frequency up to about 12 kHz where they become unpredictable due to HRTF features. Above this point and up to about 24 kHz, spectral notches exist (Slee & Young 2010). Each of these factors could contribute to narrower tuning in neurons which receive information related to the high frequency portion of the stimulus spectrum. Also, several studies have observed a negative correlation between sound

level tuning ("monotonicity") and spatial tuning area in auditory cortex, as well as a tendency for nonmonotonic units to have spatial tuning which is less expansive as level increases (Imig et al. 1990; Stecker et al. 2003; Barone et al. 1996).

First, we investigated the dependence of spatial selectivity on best frequency. In this sample, there was a small negative correlation between best frequency and tuning area ($r = -0.24$, $p < 0.001$) and a small positive correlation between best frequency and tuning vector magnitude ($r = 0.19$, $p < 0.01$). There was also a difference in the distributions of best frequency between areas, with the sample population from R/RT having a stronger representation of low frequencies and the sample population from CM/CL having a stronger representation of high frequencies (ANOVA, $p < 0.001$). It was therefore tested whether the differences in tuning width and vector magnitude between areas would be accounted for by differences in the distributions of best frequency between areas, and found that they could not (ANCOVA, area: $p < 0.05$). No effect of area on the dependency between frequency and tuning width was observed (ANCOVA, area*best frequency: $p = 0.45$). The same results were obtained for tuning vector magnitude (ANCOVA, area: $p < .001$; ANCOVA, area*best frequency: $p = 0.5$).

Next we tested the dependence of spatial selectivity on monotonicity. For this measure, an augmented measure of monotonicity (see chapter 6.1, equation 6.2) was used, as a significant portion of the sample population exhibited complex rate level tuning (Pfingst & O'Connor 1981). See chapter 6 for a discussion of rate level tuning observed in the present study. Consistent with previous results, there was a negative correlation between augmented monotonicity index (MI) and tuning area and a positive correlation between monotonicity and tuning vector magnitude (MI vs. tuning area: $r = 0.28$, $p << 0.001$; MI vs. tuning vector magnitude: $r = -0.16$, $p < 0.01$). Correlations were still observed using the classic measure of monotonicity ($R_{\max SPL}/\max(R)$), but were smaller (MI vs. tuning area: $r = 0.25$, $p << 0.001$; MI vs. tuning vector magnitude: $r = -0.14$, $p < 0.05$), suggesting that this augmented measure slightly better captured the underlying

relationship between the two neural properties. Like best frequency, an analysis of covariance suggested that monotonicity also did not account for the differences in tuning area (ANCOVA, area: $p << .001$) and vector magnitude across auditory areas (ANCOVA, area: $p << .001$). Finally, we tested whether nonmonotonic neurons showed less tuning area expansion as sound level increased. Units were classified as either "monotonic" or "nonmonotonic" based on a nonmonotonicity threshold of 0.5 and performed an analysis of covariance against sound level and monotonicity for tuning area. No significant interaction of sound level and monotonicity was found (monotonicity*SPL shift: $p = 0.4$), suggesting that spatial tuning selectivity in monotonic neurons was not less tolerant to changes in sound level.

3.1.7 Estimating spatial tuning using random spatial profile (RSP) stimuli.

Spatial receptive fields in a spatially dense acoustic environment were characterized by playing sets of broadband sounds from the entire 24-speaker array simultaneously, randomizing the sound level from each speaker. These stimuli are referred to as "random spatial profile" (RSP) stimuli. From these responses, receptive fields were constructed by calculating the linear weights (in units of spikes/second/dB SPL, see methods) rather than averaging firing rates to individual stimuli. This method is mathematically identical to a method for analyzing frequency tuning in the auditory system, using stimulus sets of random spectral shape (RSS) stimuli (Yu & Young 2000; Barbour & Wang 2003a). In these studies, receptive fields could be either similar to or quite different from those measured using pure tone stimulation.

3.1.8 Comparison of RSP weighting functions to single speaker receptive fields

Here, we quantified the similarity between the calculated weights and the single-speaker spatial receptive fields in a sample population of 77 units. A subset of these neurons showed relatively high agreement between 1- and 24-speaker conditions, but the majority of neurons had a correlation coefficient of less than 0.5 (53/77 units). Examples of units with agreement and

disagreement between conditions are shown in figure 3.14.1-6. Among those neurons showing good correlation, several showed suppressive regions which were not evident in single speaker spatial receptive field measurements. The distribution of correlation coefficients between 1 and 24 speaker conditions, as well as the angles between the tuning vectors measured with each method, are shown in figure 3.15A,B.

The spatial weighting functions generated using the responses to 24-speaker stimulation were generally not predictive of responses to such stimuli. Prediction quality was evaluated using a quality factor used previously for spectral weighting functions (Yu & Young 2000; Barbour & Wang 2003a). Quality factor (Q) is based on the mean- squared error between the predicted and observed rates divided by the variance of the predicted rates. Figure 3.15C shows the distribution of quality factor values measured in the sample population.

3.1.9 Relationship between tuning characteristics and 1- and 24- speaker receptive field correlation

In addition to the generally poor agreement of spatial receptive field shape between the 1- and 24-speaker measurements, we did not find any SRF parameters which could accurately predict 1- and 24- speaker receptive field correlation. The best predictor was the 1-speaker tuning vector magnitude, which was weakly correlated with 1- to 24-speaker agreement ($r = 0.24$, $p = 0.03$). Based on this finding, it was hypothesized that a better predictor might be the length of the tuning vector projection into the median plane. The reason for this is the fact that in multi-source listening environments, interaural properties interact more linearly in space: when coherent sounds are presented from two sources simultaneously, the resulting perceived sound source is usually between the two locations; this phenomenon is referred to as “summing localization” (Snow 1954; Leakey 1959; Blauert 1997). When sounds are played simultaneously in the median plane, however, the perceived location is that with pinna cues most closely matched to the

resulting sum of the competing sounds at the eardrums rather than at some point between the two sources (Bremen et al. 2010). Therefore, it was predicted that neurons with little selectivity in the median plane (i.e. with small median plane vector projections) should be less sensitive to spectral cues and would respond more linearly to multiple source stimulation. However, this measure was not at all predictive of 1- and 24- speaker receptive field similarity ($r = -0.014$, $p = 0.9$).

Other measures, such as tuning area ($r = -0.11$, $p = .4$), monotonicity index ($r = -0.09$; $p = 0.54$), best frequency ($r = -0.01$, $p = 0.1$), and 1-speaker tuning stability ($r = 0.21$, $p = 0.1$) did not significantly predict 1- and 24- speaker agreement. This suggests that highly directional neurons, but not all highly spatially selective neurons, may have a higher degree of spatial tuning accuracy in a multi-source auditory environment. However, considering the number of tests run, it is possible that this correlation is spurious. No effect of area on 1- and 24-speaker tuning agreement was found, but this may be due to the relatively small samples size.

3.1.10 1- and 24- speaker correlation and coding location in a noisy environment

It is possible that neurons which show relatively stable spatial tuning in single and multiple source stimulus conditions are specifically suited to encoding sound location in noisy environments. In a small number of neurons, spatial tuning with a spatially diffuse noise masker was measured. The noise masker condition was identical to the 1-speaker condition except that for each stimulus, the remaining 23 locations played noise tokens attenuated relative to the primary stimulus. In some neurons, spatial tuning in the masking condition was profoundly different from the 1-speaker SRF. In others however, the overall shape of the receptive field remained similar but was compressed in dynamic range. Only a small population was tested both with the spatial masker and the RSP stimulus ($n = 6$), but there was a significant correlation between the how well each neuron's 24-speaker weighting function correlated with the 1-speaker condition, and how well the receptive field measured in spatially diffuse noise correlated with the

1-speaker condition (@masker attenuations = 20dB; r = 0.86, p = 0.03). Several examples of spatial receptive fields measured with the spatial noise masker are shown in figure 3.16.

3.1.11 Increasing RSP stimulus tokens vs. increasing the number of stimulus repetitions: a methodological consideration

Typically, receptive fields are estimated by measuring a neuron's response to several repetitions of a relatively small stimulus set to attempt to obtain an accurate measure of the response rate to each individual stimulus. However, in this case, the goal is not to estimate the response to each individual stimulus, but rather the weighting function for the individual stimulus components. At the beginning of this study, the number of stimulus repetitions was maximized rather than the number of stimulus tokens. When preliminary results yielded low 1- and 24- stimulus agreement, it was decided to try increasing the number of stimuli to obtain better fits. This did seem to indeed increase the 1- 24- speaker correlation slightly; the mean correlation using 5 repetitions of the smaller stimulus set (51 stimuli) was 0.31 and the mean correlation of a single repetition of the larger stimulus set was 0.34, but the difference was not significant (only 10 neurons were tested with the small stimulus set). Mathematically, increasing stimulus number should increase the weight estimation accuracy (Yu 2003). Figure 3.17 summarizes the effects of increasing the number of stimuli vs. increasing the number of repetitions. From the figure, it seems that increasing the number of stimulus repetitions provides little benefit to the estimation past the first several repetitions, whereas increasing the number of stimulus sets increases the correlation value gradually and does not appear to saturate.

3.2. Discussion

3.2.1 Summary of findings

The data in this chapter represent the first measurements of full-field spatial receptive fields in an awake primate. We found that the distributions of spatial receptive fields were highly heterogeneous, with neurons tuned to contralateral and ipsilateral locations, above and below the horizon, and in the rear as well as the front. Principal component analysis corroborated this observation that variation in spatial tuning occurred along biologically relevant directions, with principal components of the distribution of receptive fields representing roughly response contrast, hemifield preference, laterality, top/down variability, and front/back variability. Of all of the directional principal components, hemifield preference accounted for the most variability across the population. There was a contralateral bias, which was strongest in the caudal areas CM/CL. Despite this bias, a significant portion of neurons were tuned to the ipsilateral hemifield. SRF shapes were found to vary in a continuous manner, as we did not find evidence for clusters of receptive field shapes. The majority of neurons, particularly in areas A1 and CM/CL had relatively narrow tuning areas, with responses greater than half the maximum firing rate frequently encompassing less than 25% of space. A complementary measure of directionality, the tuning vector magnitude, which is sensitive to dispersion of spatial receptive fields, showed that as a population, directionality is not very high. This is due to the fact that neurons can have non-contiguous receptive fields, or be narrowly tuned in only a single dimension. Tuning area, vector, and vector magnitude were not highly sensitive to changes in sound level. Lastly, a novel approach was introduced to characterize spatial tuning in a multi-source acoustic environment by presenting sounds from all sound sources simultaneously at randomized levels. Using this method it was found that most neurons exhibit drastically different spatial tuning in a multi-source environment than in the single source condition, probably reflecting highly nonlinear mechanisms underlying spatial processing.

3.2.2 Comparison with previous studies: effects of anesthesia

To our knowledge there are no studies of full-field spatial tuning in awake mammals. However, we can compare the present results with studies conducted in anesthetized animals and studies of azimuth tuning. The fact that the two groups are not mutually exclusive may provide insight into potential factors underlying the differences between these data and those reported previously. Studies of spatial tuning in anesthetized animals (Middlebrooks et al. 1994; Brugge et al. 1996; Stecker et al. 2003; Mrsic-flogel et al. 2005), have observed receptive fields which, at low sound levels, are typically centered in the contralateral-frontal quadrant of space, along the "acoustic-axis" of the ear. As sound levels increase, these receptive fields broaden, often dramatically. This type of spatial response is particularly prevalent in studies of full-field spatial tuning in cats (Brugge et al. 1994; Brugge et al. 1996) and ferrets (Mrsic-flogel et al. 2005), carried out in anesthetized animals. These findings are in contrast to the relatively stable behavioral acuity for spatial localization across sound levels. More recent studies in awake animals have found very small effects, if any, of sound level on spatial tuning area. These studies, however, have focused either on sounds varying in the azimuth dimension (Mickey & Middlebrooks 2003; Woods et al. 2006) or a portion of the frontal hemifield (Recanzone et al. 2000; Zhou & Wang 2012). It was shown here that the stability of receptive field areas are not an result of measurement of only azimuth locations, with both tuning area and tuning vector magnitude staying stable across sound levels within single neurons. In fact, neurons in CM/CL had a tendency to decrease tuning area with increasing sound levels. We did notice that in A1 and R/T, but not in CM/CL, population tuning area tended to increase slightly with increasing sound level. Previous studies have suggested that neurons' level tuning functions may be linked with spatial tuning over level. Specifically, it has been observed in anesthetized animals that neurons with nonmonotonic rate level functions show more spatial selectivity and also less receptive field expansion with increasing sound levels (Imig et al. 1990; Barone et al. 1996; Stecker et al. 2003). Although

nonmonotonic neurons in this study did show a higher degree of spatial selectivity, monotonic neurons were not found to have significantly expanding spatial receptive fields with increasing sound level.

The stark differences between the observed effects of sound level on spatial selectivity in studies of anesthetized and awake animals underscores the second major difference between present results and previous full field studies: the broad distribution of spatial receptive fields. Although previous studies of full-field tuning have found some variation in receptive field shape, responses across the populations were strongly centered in the frontal-contralateral hemifield. This has led some to hypothesize a decoding scheme in which azimuth can be decoded by subtracting responses from each cortical hemisphere (Stecker et al. 2005). In contrast, we found neurons with roughly equal distributions in the front and rear as well as the top and bottom hemifields. Previous studies of frequency tuning in awake animals have shown that suppressive and nonlinear interactions may play a large role in shaping responses in awake animals (Barbour & Wang 2003a; Kadia & Wang 2003; Sadagopan & Wang 2009). In contrast, spatial receptive fields in A1 in anesthetized animals have been shown to be largely explained by a linear superposition of inputs across frequency and binaural inputs (Schnupp et al. 2001). It is likely that spatial receptive fields in awake animals may be governed by these same nonlinear interactions seen in studies of frequency integration. Closed field studies in awake animals are needed to further investigate this possibility.

3.2.3 Comparison with previous studies: species specific differences

It is also possible that there exist species specific differences in spatial tuning properties in auditory cortex. In cats, several groups have reported receptive fields which typically span an entire hemifield, with many neurons responding omnidirectionally (Brugge et al. 1996; Stecker et al. 2003), leading to a hypothesis for a "panoramic" code for auditory space (Middlebrooks et al.

1994). This preponderance of broadly tuned receptive fields may persist in awake cats; median tuning areas as high as 294° have been measured in primary auditory cortex (Lee & Middlebrooks 2011), which is equivalent to a tuning area of about 0.8. In anesthetized ferrets, spatial receptive fields rarely exceeded a hemifield in size, although these estimates may be biased towards low tuning areas by a high degree of noise in firing rate estimates due to neuronal noise as a result of low number of stimulus repetitions, a tradeoff of high density spatial sampling (Mrsic-flogel et al. 2005). In a study of awake macaques, median tuning areas ranged from ~90° to ~270° in A1 (0.25 to 0.75) to ~90° (0.25) in CL, with which the present azimuth tuning area means of .45 in A1 and .31 in CM/CL are roughly in agreement (Woods et al. 2006). Finally, a recent study in anesthetized rats found that neurons in auditory cortex were exclusively tuned to the contralateral hemifield, an observation that to our knowledge has not been made in other species (Yao et al. 2013). Additionally, spatial receptive field areas were highly tolerant to increasing sound level, which again to our knowledge has not been observed in an anesthetized animal. Thus it may be possible that spatial coding strategies may vary among mammals from species to species, with primates favoring a more localized, rather than panoramic code.

3.2.4 Spatial tuning in primary, rostral, and caudal areas

We found that, in comparison with the primary auditory cortex (A1), the caudal areas CM/CL had sharper spatial receptive fields, whereas those in the rostral areas R/RT were broader. These results are consistent with previous studies in several species, both in anesthetized and awake animals (Rauschecker & Tian 2000; Tian et al. 2001; Stecker et al. 2003; Woods et al. 2006). It has been suggested that this increase in spatial selectivity in the caudal areas may be related to an increased level of nonmonotonicity of rate-level tuning. Here, more nonmonotonic neurons did tend to have narrower spatial receptive fields, but that this difference could not be explained by differences in rate level tuning between areas. It was also observed that spatial receptive fields were more contralaterally biased in CM/CL, although normalized population firing rates were

roughly flat within the contralateral hemifield in all areas. This increase in firing rate also translated to an increase in population firing rate separation between different locations, measured by Euclidean distance, in the contralateral hemifield. It is noted, however, that this increase came at the expense of population distance in the ipsilateral hemifield, and that distances in A1 and even R/RT were not categorically smaller. In other words, there exists significant spatial information in all areas measured. This is underscored by the fact that receptive field shapes varied continuously within and across areas, as attempts to cluster and classify neurons had little success.

3.2.5 Coding space in a spatially dense acoustic environment

Borrowing mathematical concepts from previous studies of spectral integration in subcortical (Yu & Young 2000; Young & Calhoun 2005) and cortical (Barbour & Wang 2003a) areas, we investigated coding for spatial location in auditory cortex in a spatially dense acoustic environment. While many neurons responded strongly to these random spatial profile (RSP) stimuli, the majority of weighting functions calculated using these responses bore little resemblance to receptive fields measured using single-source sound stimulation. This is perhaps not surprising considering the nonlinear mechanisms underlying the computation of sound source location in the brain (Goldberg & Brown 1969; Yu & Young 2000; Peña & Konishi 2001). We hypothesized that neurons which displayed little median plane sensitivity would show more linear RSP responses and therefore higher 1- and 24- speaker correlation, as sound localization in the median plane is nonlinearly degraded by the existence of competing sound sources (Bremen et al. 2010). However, median plane selectivity did not predict 1- and 24- speaker correlation. Several other measures were tested, including tuning area, tuning vector magnitude, best frequency, and monotonicity. Only one of these measures, tuning vector magnitude, showed a significant correlation with 1- and 24- speaker correlation, and it also was a relatively poor predictor.

Some units, however, did show high 1- and 24-speaker agreement. We hypothesized that such neurons may be well-suited to encoding spatial location in noisy environments. In a small number of neurons, spatial receptive fields were measured with a spatially diffuse noise masker, a good agreement between was found between neurons' tolerance to spatial noise and their 1- and 24- speaker agreement. Although preliminary, this finding may suggest that there exists a population of neurons in auditory cortex which is particularly well suited to encoding spatial location in noisy environments, perhaps even useful for spatial stream segregation (such as in the "cocktail party"), and that RSP stimuli are appropriate for the characterization of such neurons. Another theory, recently proposed, is that stream segregation may be accomplished by separate populations of cortical neurons synchronizing to temporal acoustic features of alternately presented sounds (Middlebrooks & Bremen 2013). It's possible that multiple mechanisms are involved in such situations.

3.2.6 Conclusions and implications for spatial and non-spatial processing

We showed that in several areas of auditory cortex, large proportions of neurons respond to a relatively restricted area of space at all sound intensities, and that these neurons' receptive fields were broadly distributed throughout space. This suggests an alternate hypothesis to the panoramic view of spatial encoding in primates: one where each location, at least within the contralateral hemifield, is equally represented by neurons. The fact that receptive field shapes are often oddly shaped, and even multi-peaked, however, would preclude the ability to employ a population vector decoding scheme. It is possible that a more generalized decoder which does not make assumptions about receptive field shapes, but does make assumptions about population firing rate distributions, such as a log-likelihood decoder (Jazayeri & Movshon 2006), is a potential candidate. Finally, the large number of narrow receptive fields could present a problem for the auditory system when trying to maintain a stable percept in a dynamic spatial environment, whether due to self or external movement. For the marmoset in particular, a highly mobile,

arboreal species, single sound sources may be represented by completely non-overlapping neural populations within a single stimulus period. The visual system encounters a similar problem when making saccades to survey a visual scene (Melcher & Colby 2008). How the visual system manages to maintain a stable percept when the image on the retina is rapidly changing is an active topic of investigation. Understanding how the auditory system binds auditory objects in a changing and sometimes unpredictable spatial environment will require recording neural activity from free moving subjects. It is likely that many brain areas, including those involved in visual, motor, and vestibular processing may be involved in accomplishing this task.

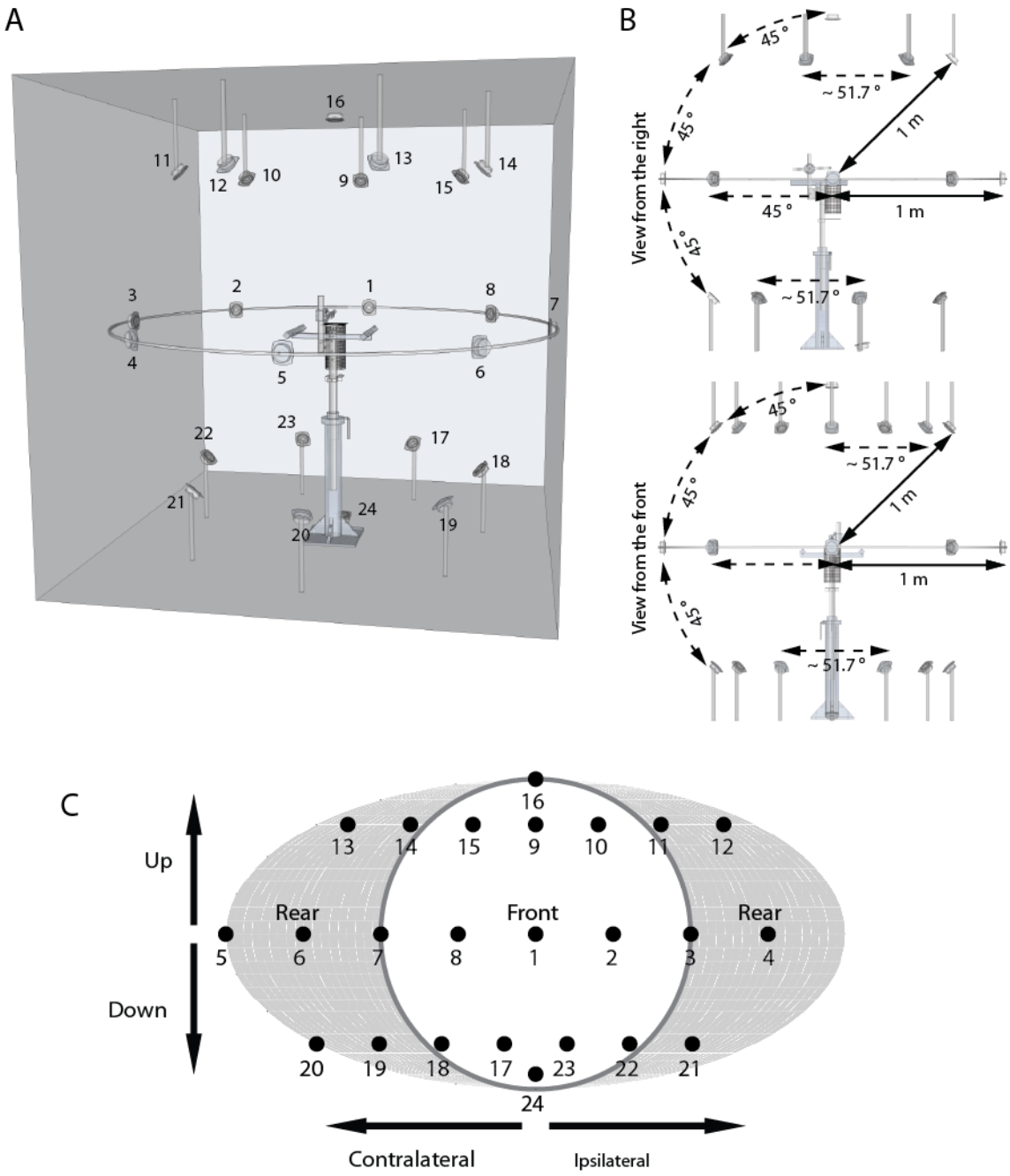


Figure 3.1. **Full free-field speaker layout.** (**a**) 3D model of sound isolating chamber with 24-speaker array, marmoset chair and single unit recording frame. This experimental configuration was designed to allow sound delivery from all directions. (**b**) Side and front views of experimental setup showing layout measurements. (**c**) A Fournier projection of space, opened at the rear meridian. As recordings were made in the left hemispheres of all subjects, left and right hemifields represent contralateral and ipsilateral space, respectively.

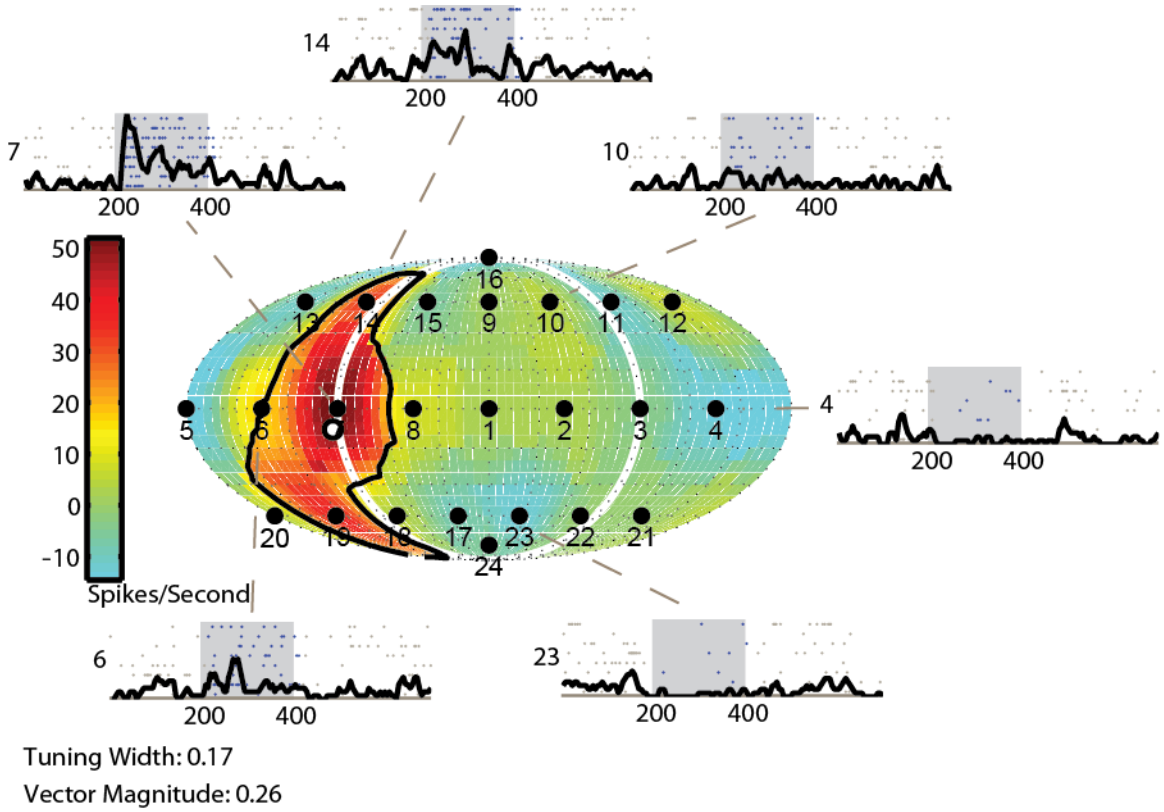


Figure 3.2. Example contralateral-tuned spatial receptive field. Peri-stimulus time histograms and spike raster plots show the time course of responses.

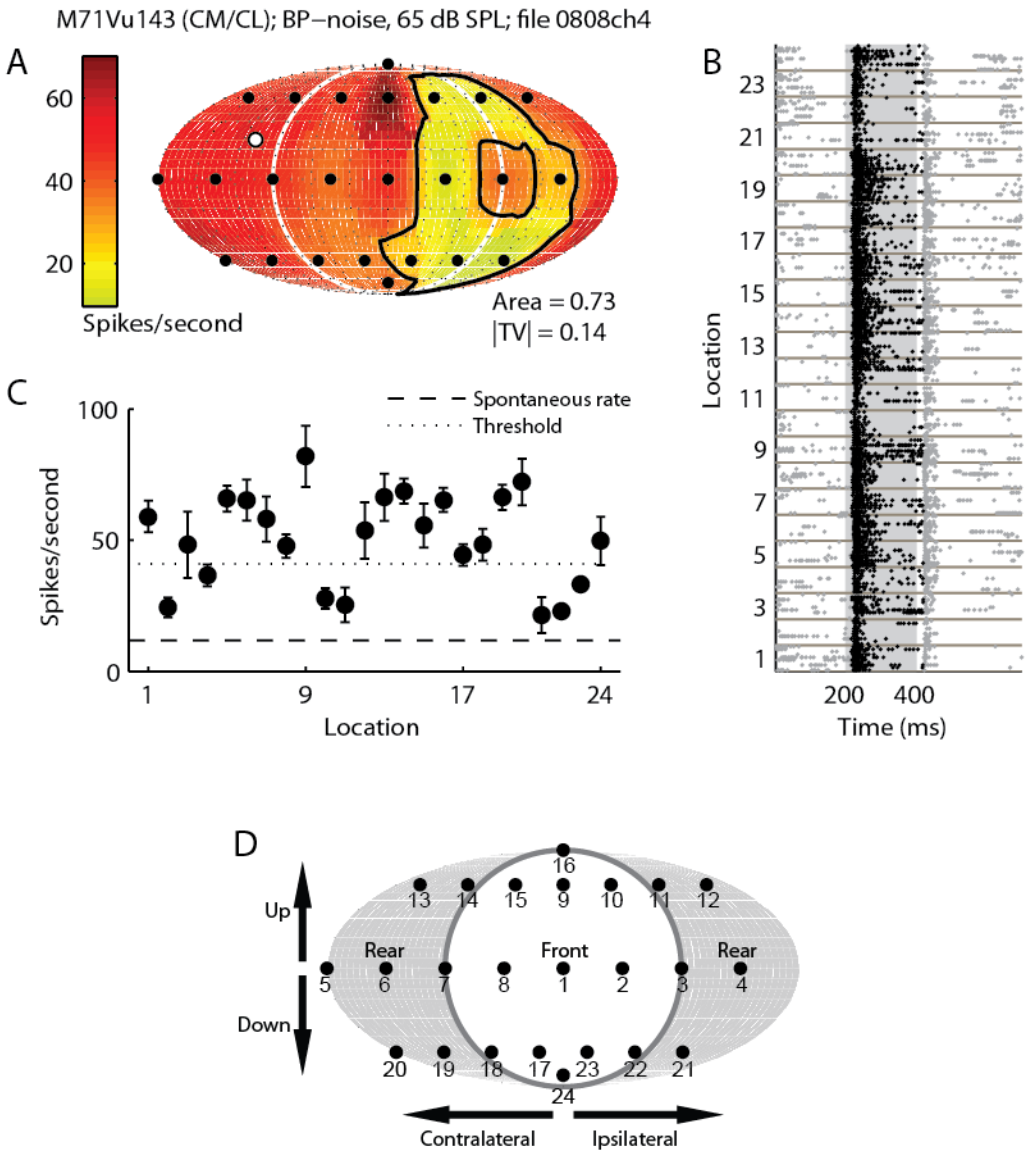


Figure 3.3.1. Spatial receptive field examples (1). Example unit responding to the entire contralateral hemifield. (**a**) Relative firing rates displayed on a Fournier projection of space. (**b**) Peristimulus time histogram (PSTH) raster plot. Analysis windows are 15 milliseconds post-stimulus onset to 20 milliseconds post-stimulus offset. (**c**) Firing rates at different speaker locations. Error bars represent standard error of the mean. The threshold is the 50% point between 0 and the maximum firing rate. (**d**) Speaker layout reference.

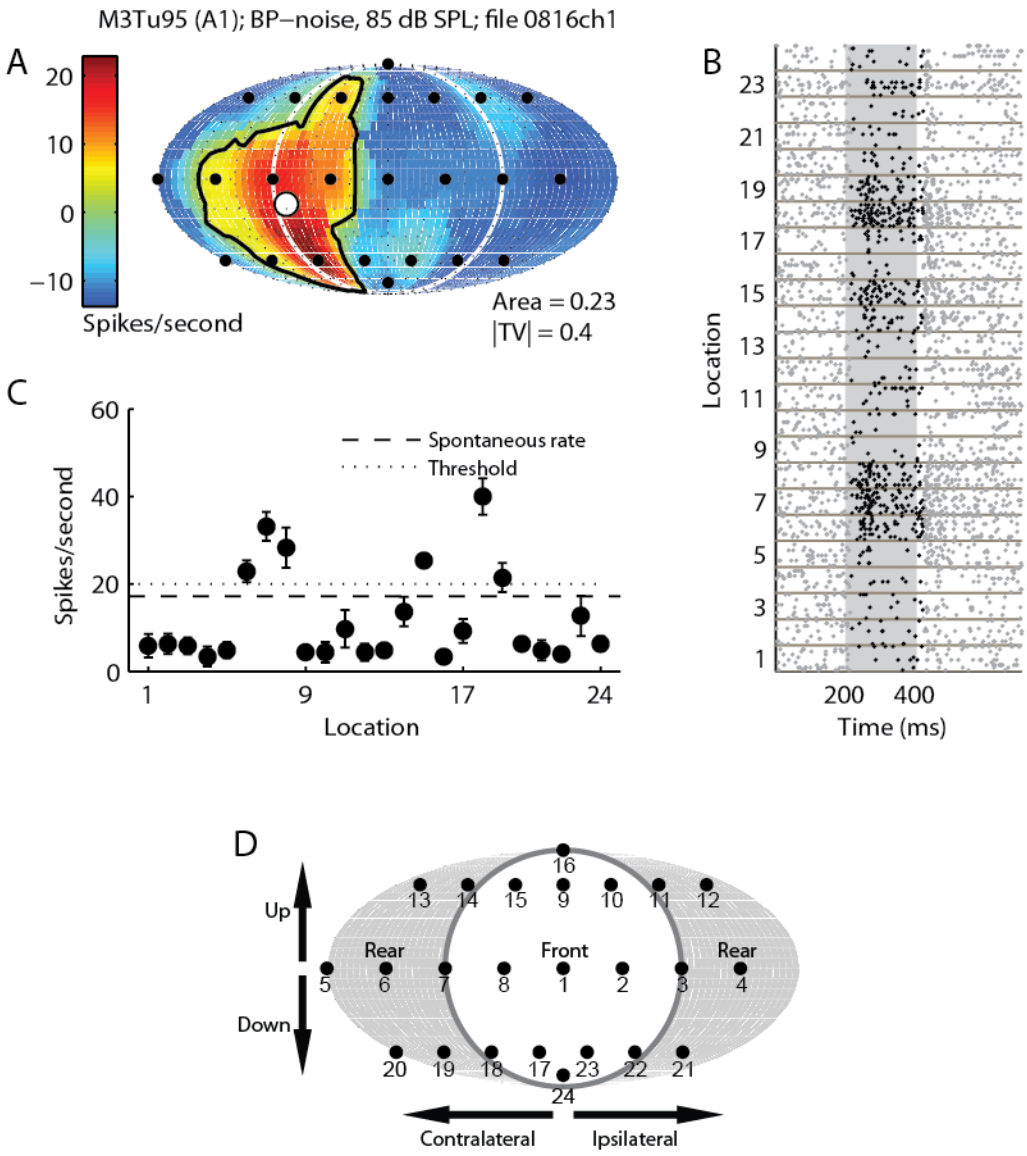


Figure 3.3.2. Spatial receptive field examples (2). Example unit responding to selectively to the contralateral hemifield hemifield. (**a**) Relative firing rates displayed on a Fournier projection of space. (**b**) Peristimulus time histogram (PSTH) raster plot. Analysis windows are 15 milliseconds post-stimulus onset to 20 milliseconds post-stimulus offset. (**c**) Firing rates at different speaker locations. Error bars represent standard error of the mean. The threshold is the 50% point between 0 and the maximum firing rate. (**d**) Speaker layout reference.

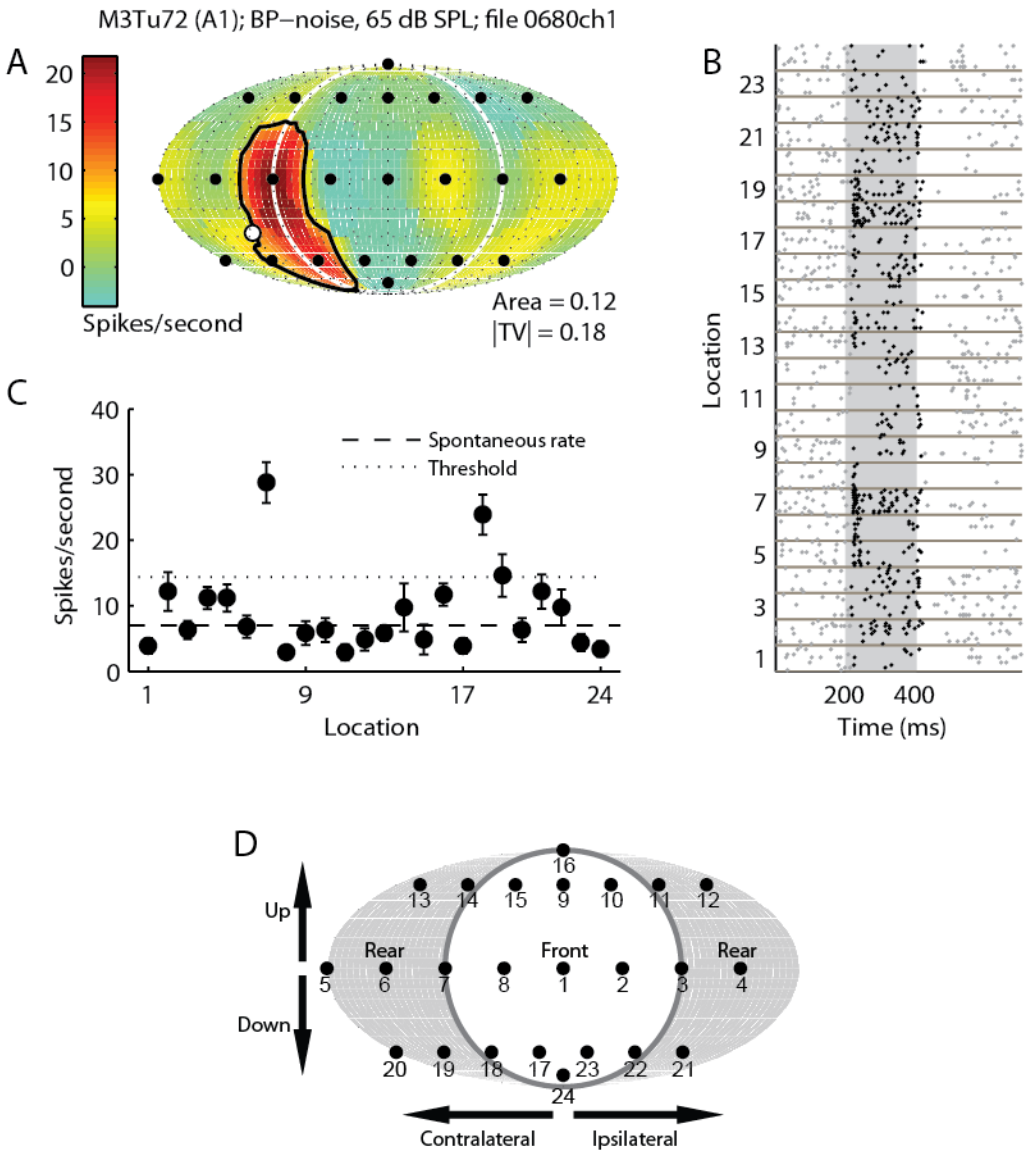


Figure 3.3.3. Spatial receptive field examples (3). Example unit narrowly tuned to the most contralateral locations. (**a**) Relative firing rates displayed on a Fournier projection of space. (**b**) Peristimulus time histogram (PSTH) raster plot. Analysis windows are 15 milliseconds post-stimulus onset to 20 milliseconds post-stimulus offset. (**c**) Firing rates at different speaker locations. Error bars represent standard error of the mean. The threshold is the 50% point between 0 and the maximum firing rate. (**d**) Speaker layout reference.

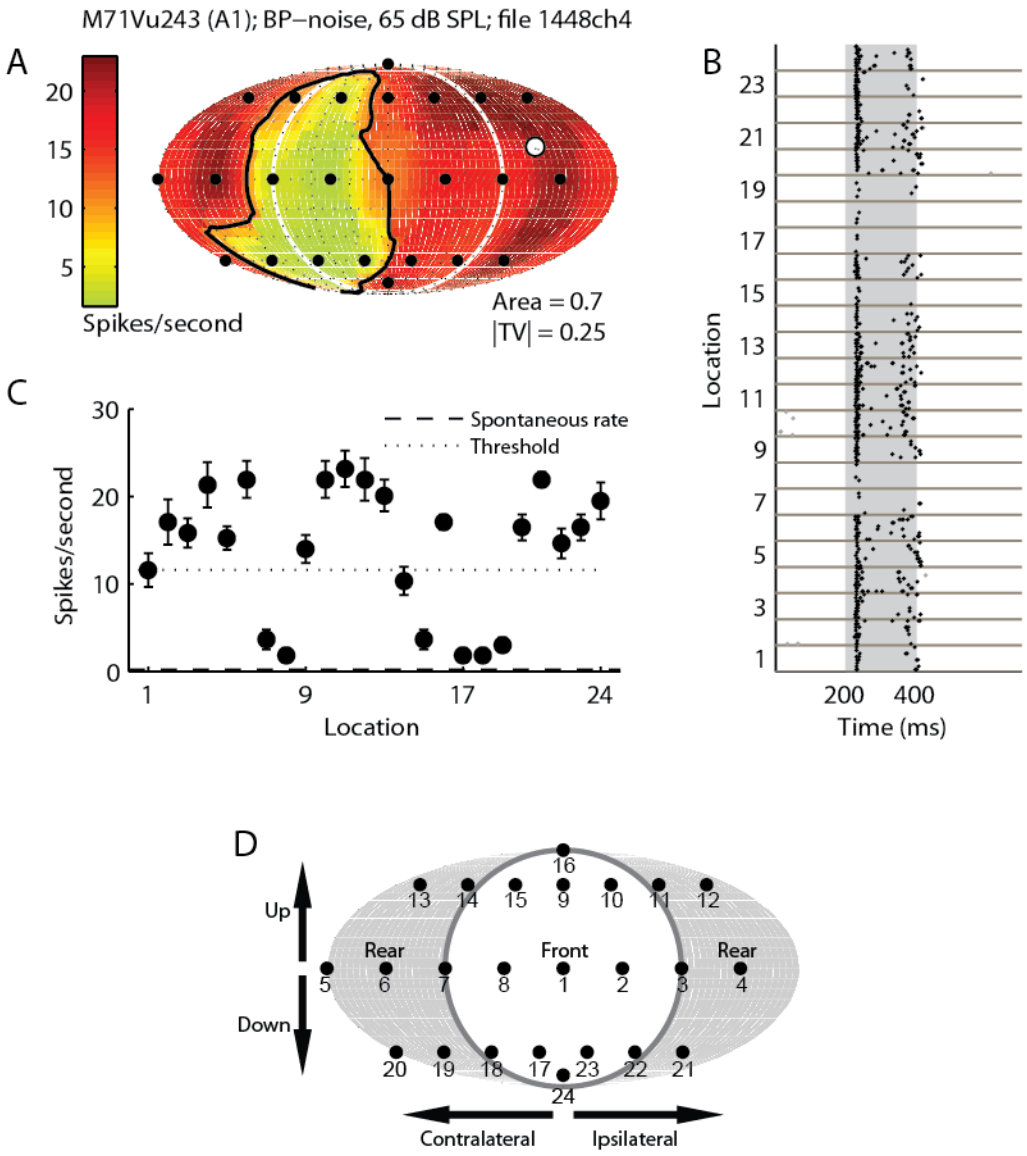


Figure 3.3.4. Spatial receptive field examples (4). Example unit broadly tuned to the ipsilateral hemifield. (**a**) Relative firing rates displayed on a Fournier projection of space. (**b**) Peristimulus time histogram (PSTH) raster plot. Analysis windows are 15 milliseconds post-stimulus onset to 20 milliseconds post-stimulus offset. (**c**) Firing rates at different speaker locations. Error bars represent standard error of the mean. The threshold is the 50% point between 0 and the maximum firing rate. (**d**) Speaker layout reference.

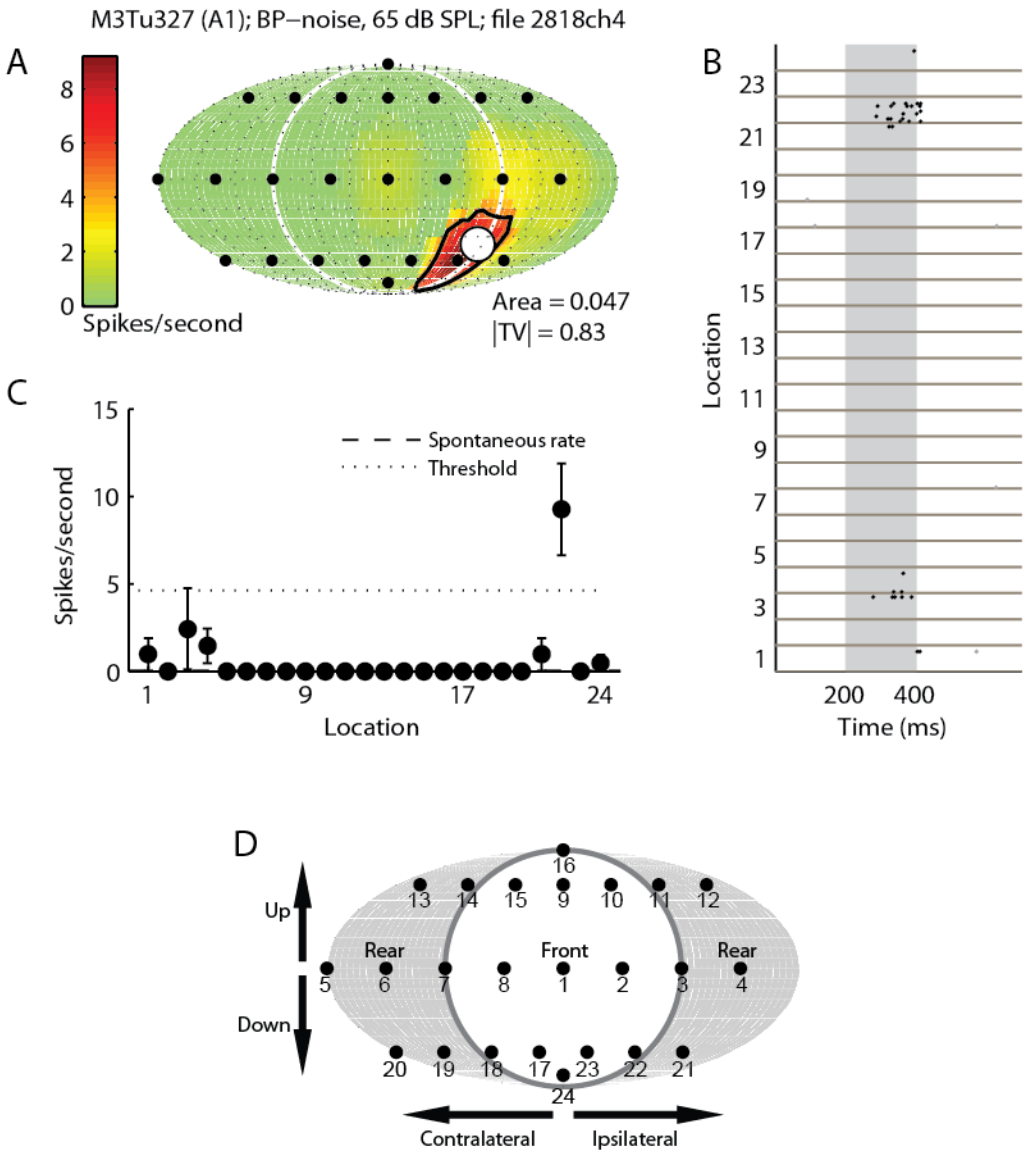


Figure 3.3.5. Spatial receptive field examples (5). Example narrowly tuned ipsilaterally. (**a**) Relative firing rates displayed on a Fournier projection of space. (**b**) Peristimulus time histogram (PSTH) raster plot. Analysis windows are 15 milliseconds post-stimulus onset to 20 milliseconds post-stimulus offset. (**c**) Firing rates at different speaker locations. Error bars represent standard error of the mean. The threshold is the 50% point between 0 and the maximum firing rate. (**d**) Speaker layout reference.

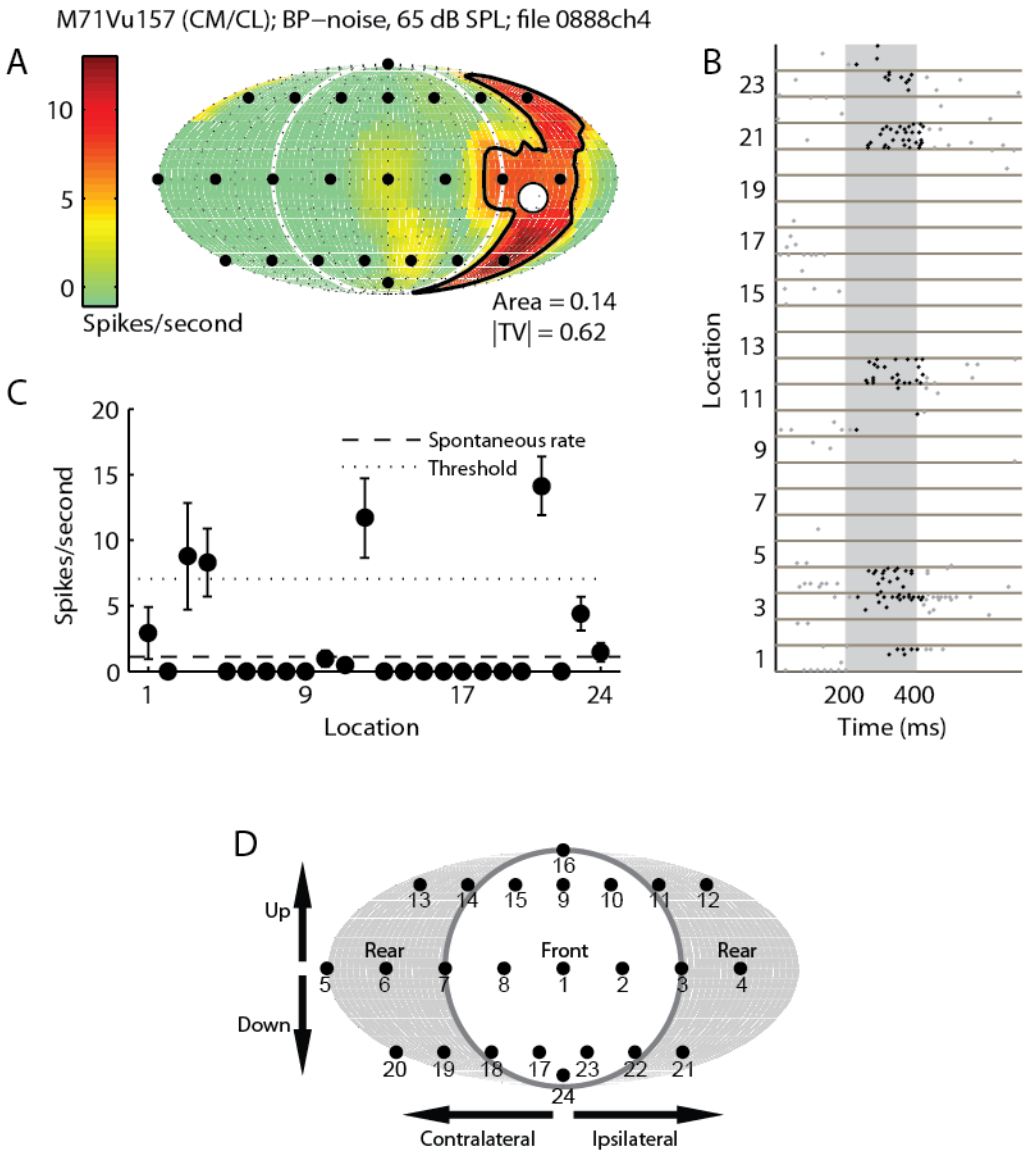


Figure 3.3.6. Spatial receptive field examples (6). Example unit tuned narrowly to ipsilateral locations. (**a**) Relative firing rates displayed on a Fournier projection of space. (**b**) Peristimulus time histogram (PSTH) raster plot. Analysis windows are 15 milliseconds post-stimulus onset to 20 milliseconds post-stimulus offset. (**c**) Firing rates at different speaker locations. Error bars represent standard error of the mean. The threshold is the 50% point between 0 and the maximum firing rate. (**d**) Speaker layout reference.

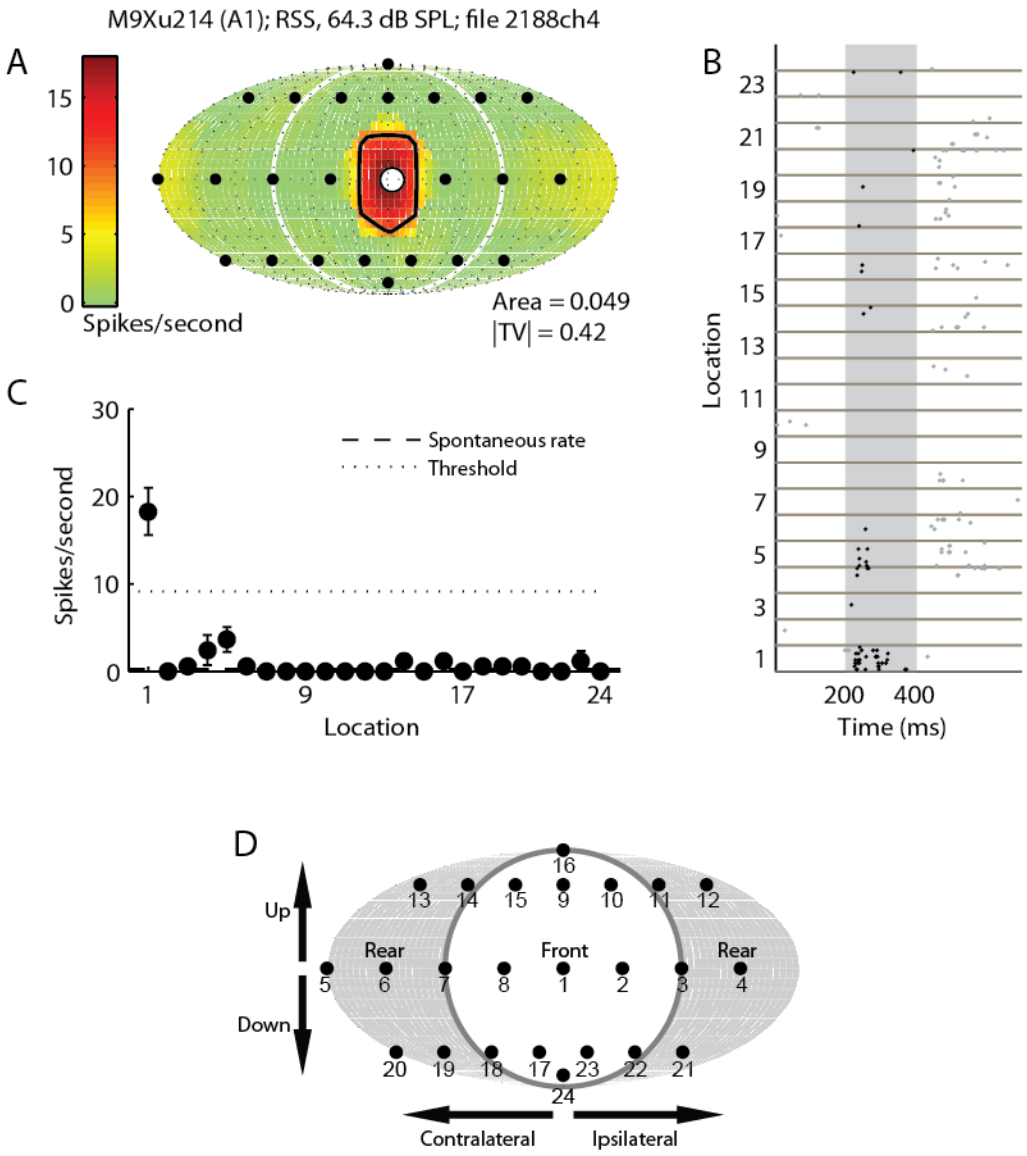


Figure 3.3.7. Spatial receptive field examples (7). Example unit tuned narrowly to the front location. (**a**) Relative firing rates displayed on a Fournier projection of space. (**b**) Peristimulus time histogram (PSTH) raster plot. Analysis windows are 15 milliseconds post-stimulus onset to 20 milliseconds post-stimulus offset. (**c**) Firing rates at different speaker locations. Error bars represent standard error of the mean. The threshold is the 50% point between 0 and the maximum firing rate. (**d**) Speaker layout reference.

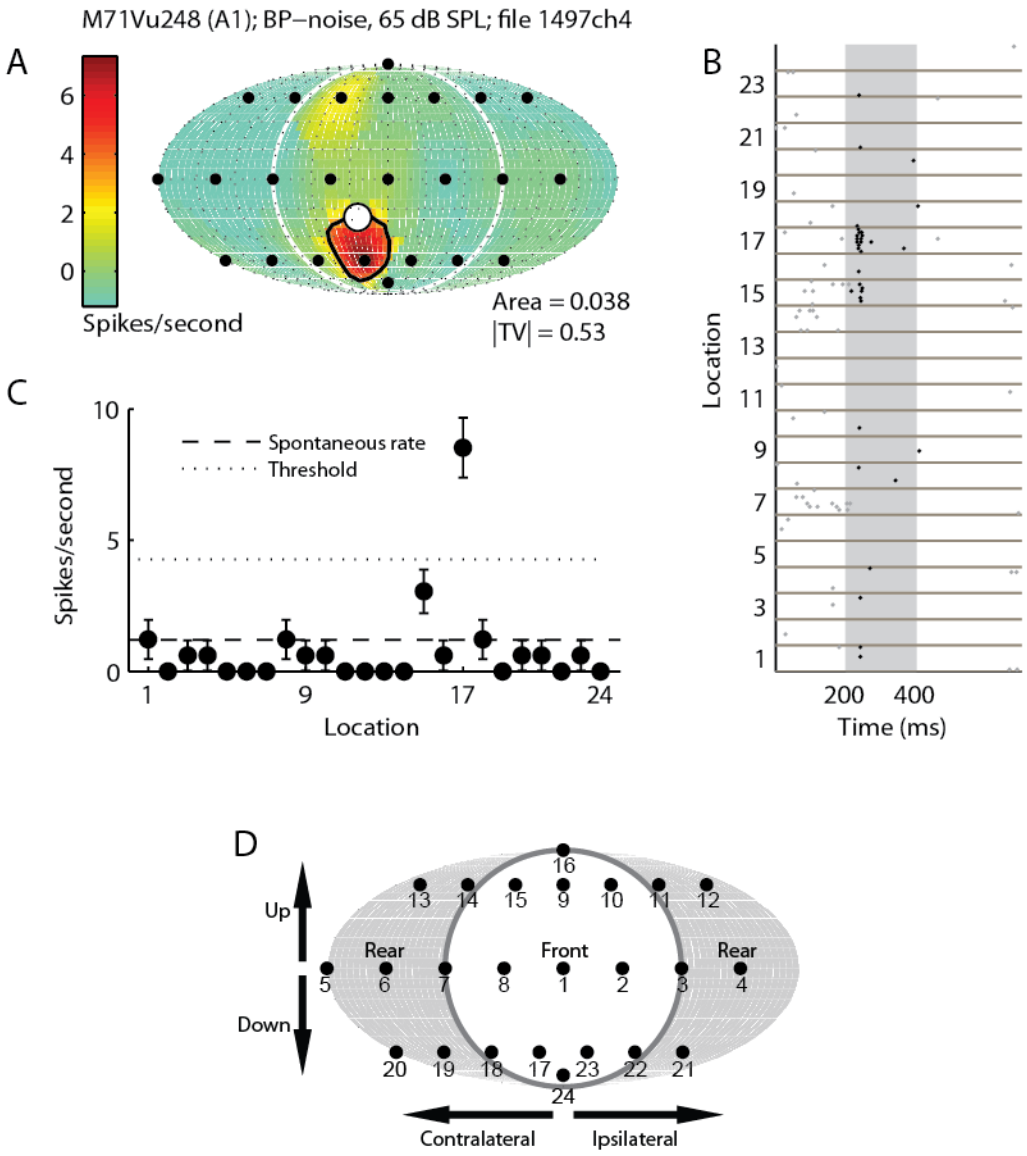


Figure 3.3.8. Spatial receptive field examples (8). Example unit (2) tuned narrowly to the front. (**a**) Relative firing rates displayed on a Fournier projection of space. (**b**) Peristimulus time histogram (PSTH) raster plot. Analysis windows are 15 milliseconds post-stimulus onset to 20 milliseconds post-stimulus offset. (**c**) Firing rates at different speaker locations. Error bars represent standard error of the mean. The threshold is the 50% point between 0 and the maximum firing rate. (**d**) Speaker layout reference.

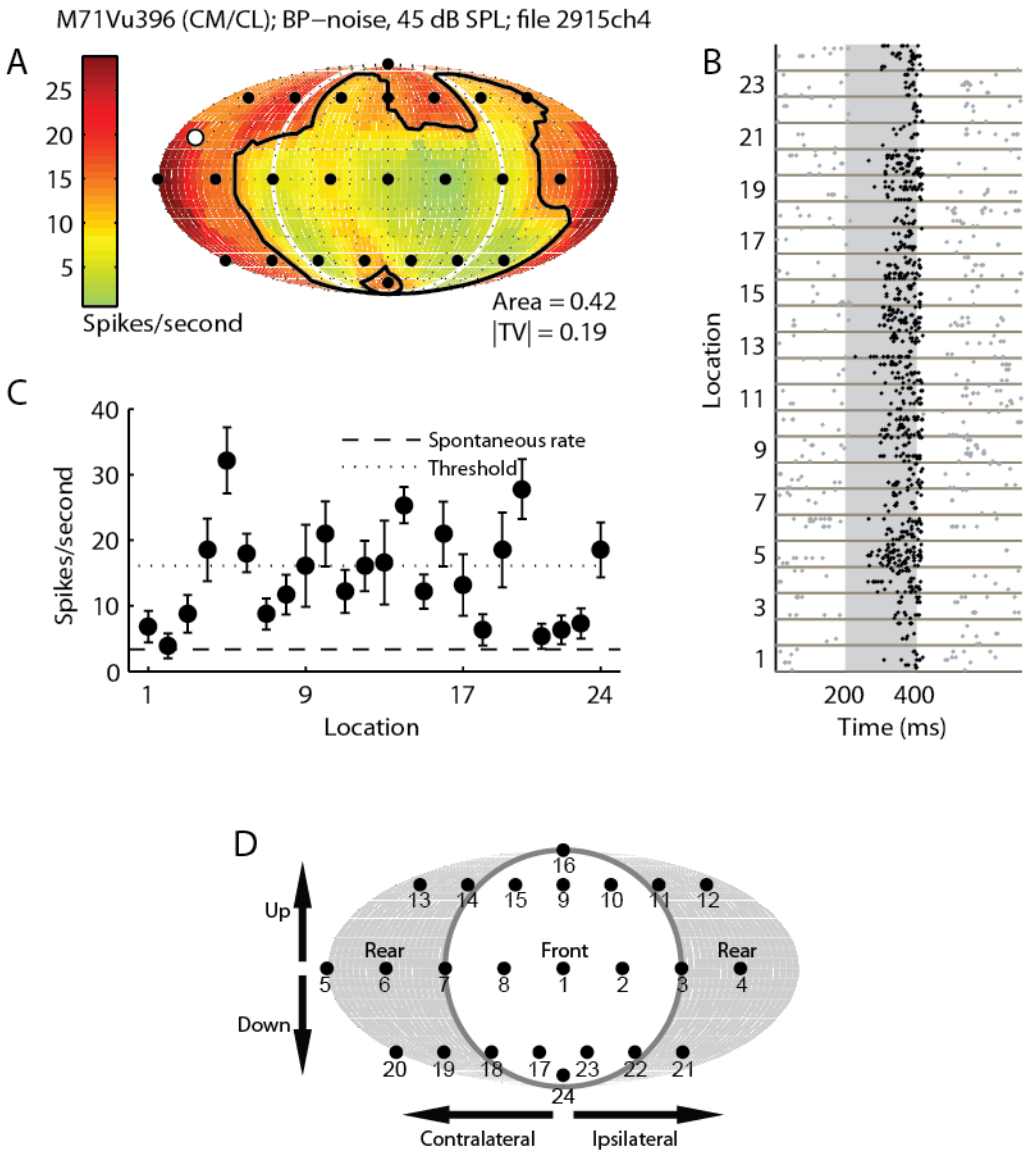


Figure 3.3.9. Spatial receptive field examples (9). Example unit tuned to the rear. (**a**) Relative firing rates displayed on a Fournier projection of space. (**b**) Peristimulus time histogram (PSTH) raster plot. Analysis windows are 15 milliseconds post-stimulus onset to 20 milliseconds post-stimulus offset. (**c**) Firing rates at different speaker locations. Error bars represent standard error of the mean. The threshold is the 50% point between 0 and the maximum firing rate. (**d**) Speaker layout reference.

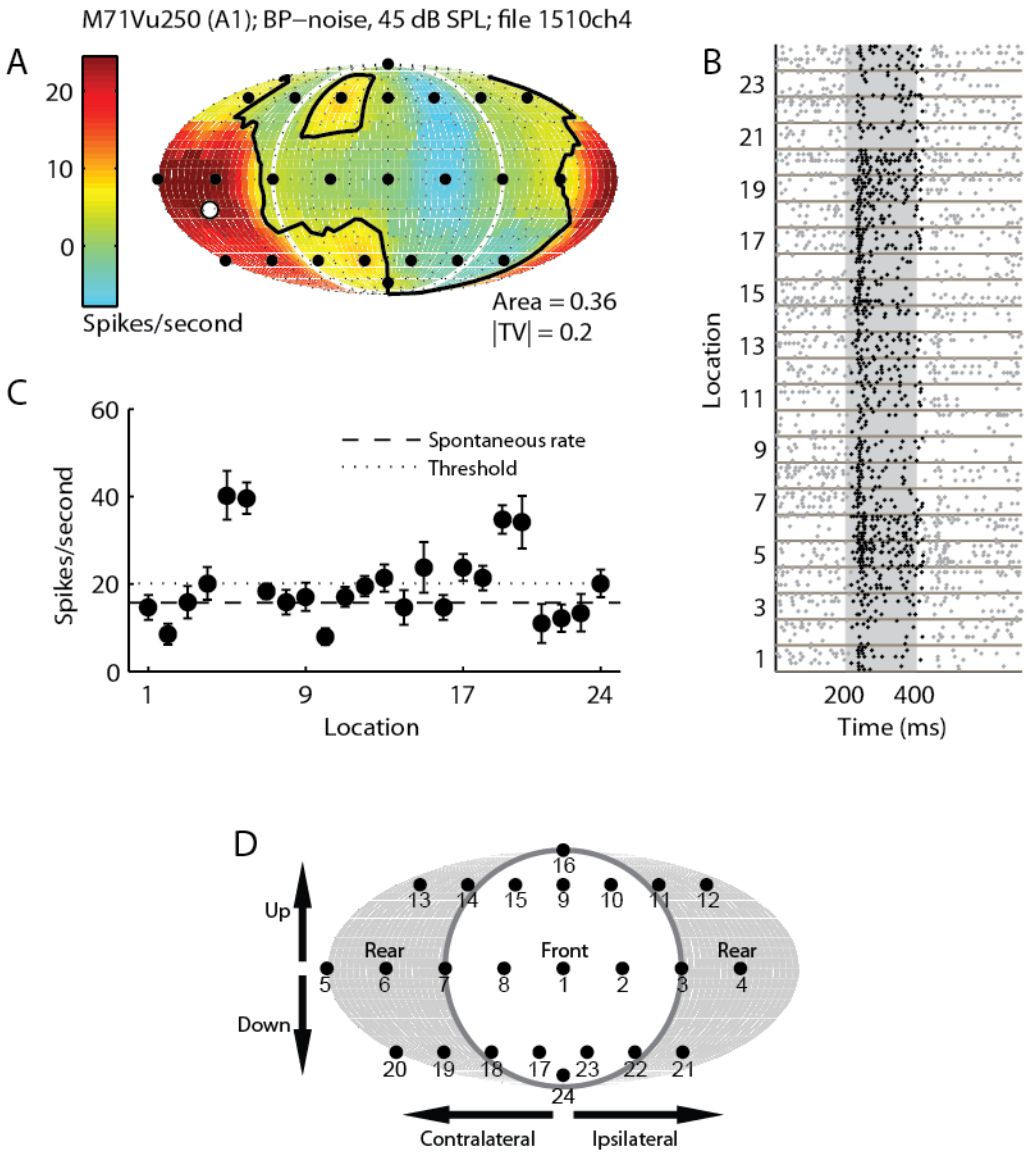


Figure 3.3.10. Spatial receptive field examples (10). Example unit (2) tuned to the rear. (**a**) Relative firing rates displayed on a Fournier projection of space. (**b**) Peristimulus time histogram (PSTH) raster plot. Analysis windows are 15 milliseconds post-stimulus onset to 20 milliseconds post-stimulus offset. (**c**) Firing rates at different speaker locations. Error bars represent standard error of the mean. The threshold is the 50% point between 0 and the maximum firing rate. (**d**) Speaker layout reference.

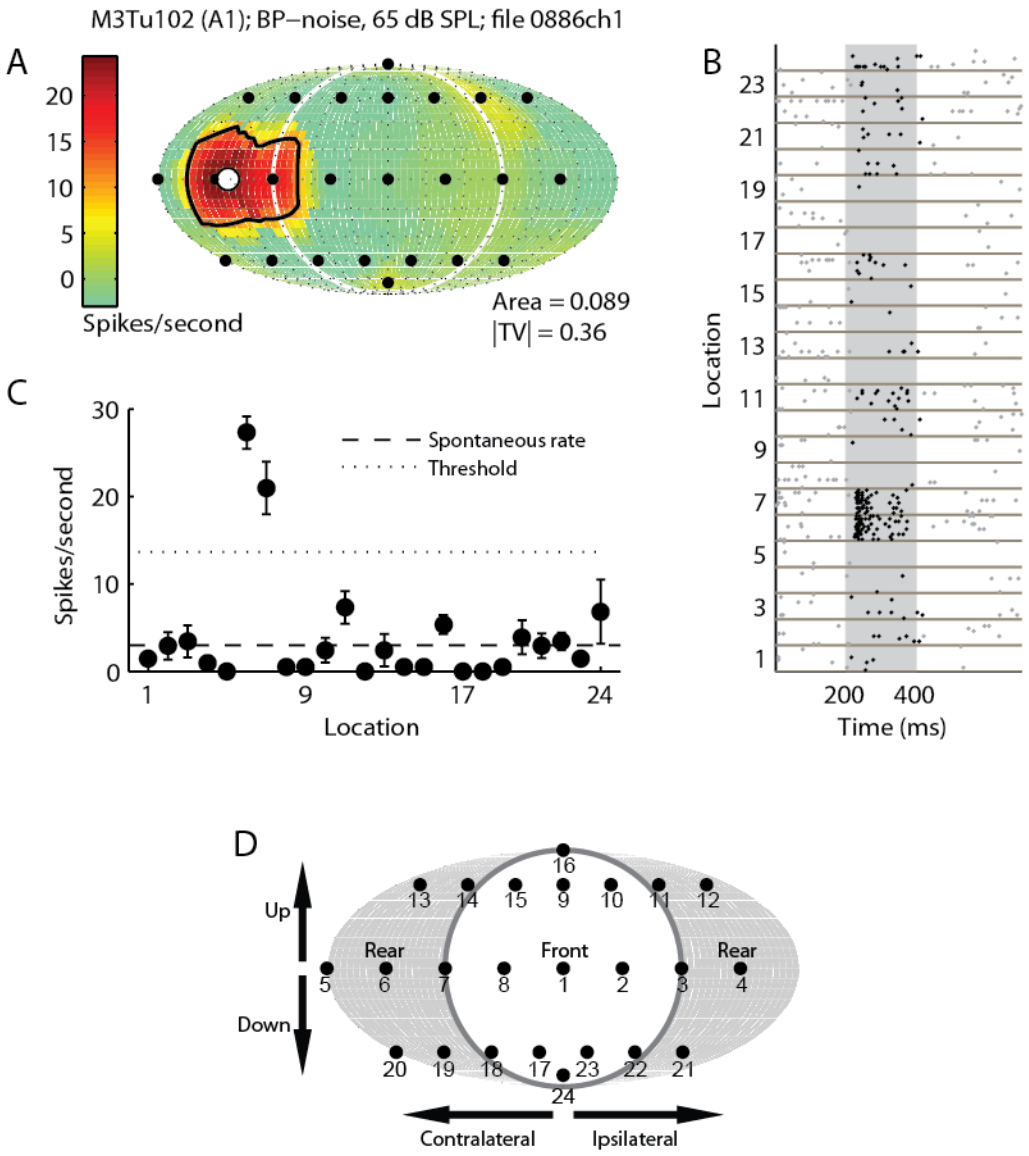


Figure 3.3.11. Spatial receptive field examples (11). Example unit (3) tuned to the rear. (**a**) Relative firing rates displayed on a Fournier projection of space. (**b**) Peristimulus time histogram (PSTH) raster plot. Analysis windows are 15 milliseconds post-stimulus onset to 20 milliseconds post-stimulus offset. (**c**) Firing rates at different speaker locations. Error bars represent standard error of the mean. The threshold is the 50% point between 0 and the maximum firing rate. (**d**) Speaker layout reference.

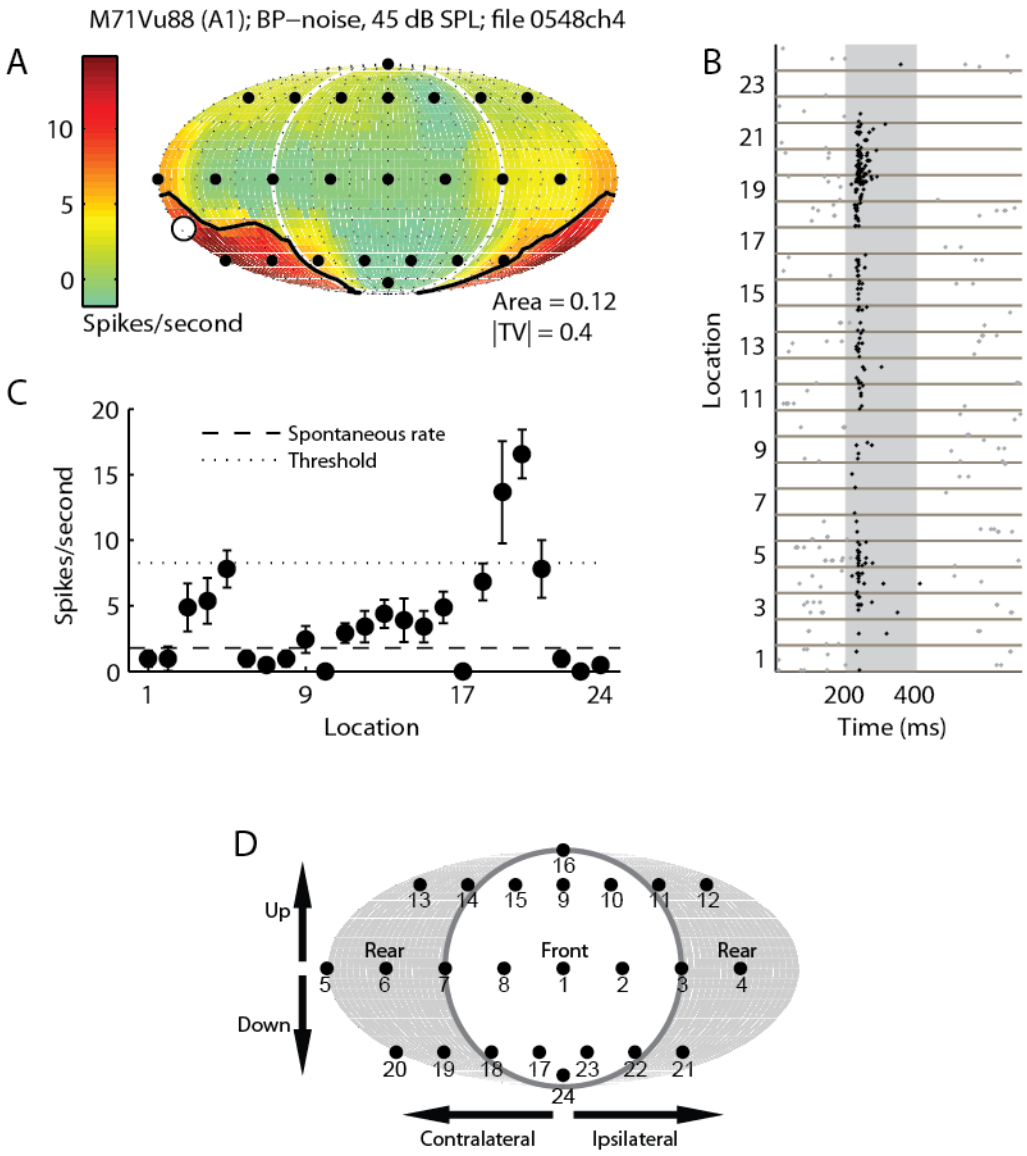


Figure 3.3.12. Spatial receptive field examples (12). Example unit (4) tuned to the rear. (**a**) Relative firing rates displayed on a Fournier projection of space. (**b**) Peristimulus time histogram (PSTH) raster plot. Analysis windows are 15 milliseconds post-stimulus onset to 20 milliseconds post-stimulus offset. (**c**) Firing rates at different speaker locations. Error bars represent standard error of the mean. The threshold is the 50% point between 0 and the maximum firing rate. (**d**) Speaker layout reference.

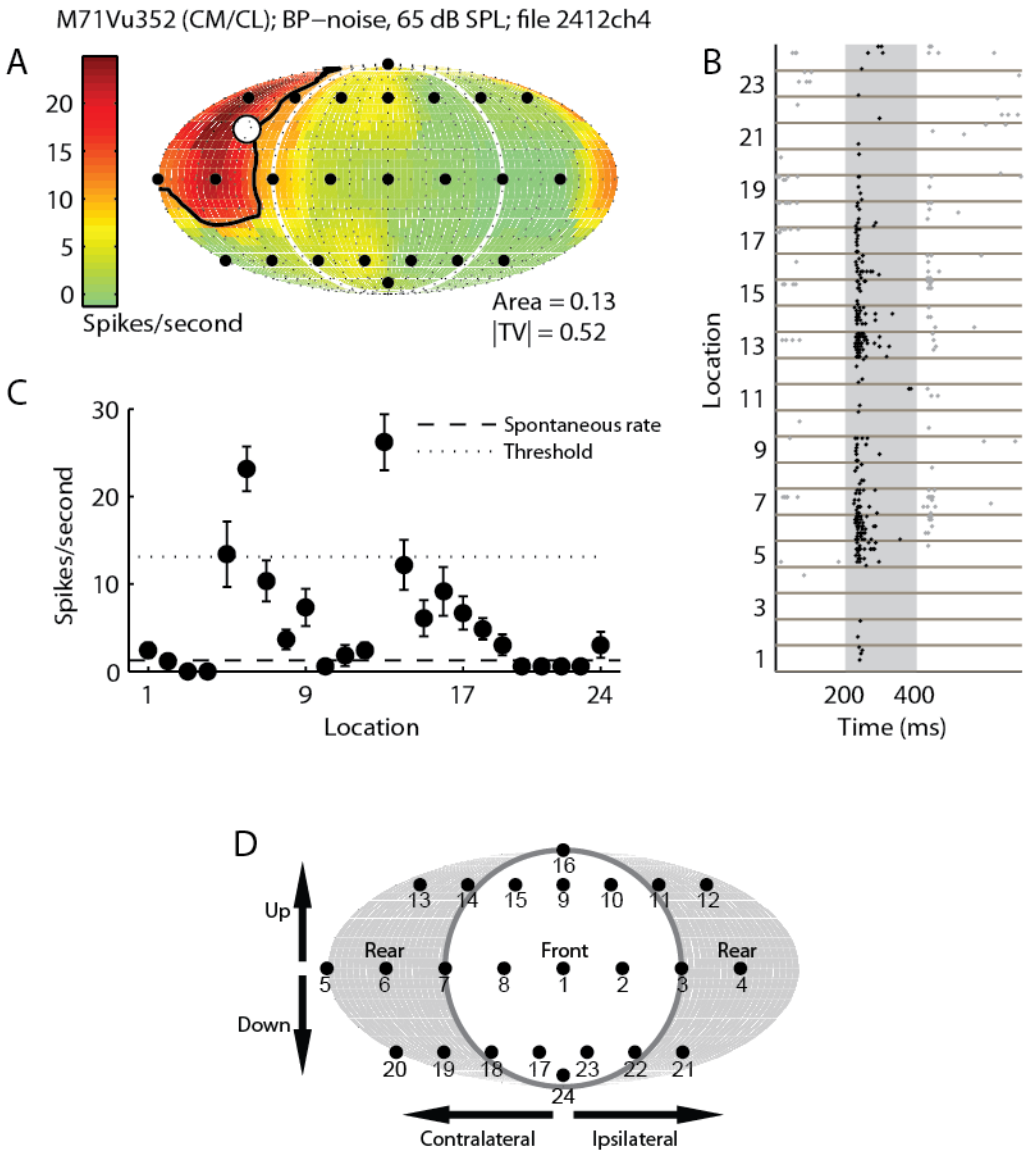


Figure 3.3.13. Spatial receptive field examples (13). Example unit (5) tuned to the rear. (**a**) Relative firing rates displayed on a Fournier projection of space. (**b**) Peristimulus time histogram (PSTH) raster plot. Analysis windows are 15 milliseconds post-stimulus onset to 20 milliseconds post-stimulus offset. (**c**) Firing rates at different speaker locations. Error bars represent standard error of the mean. The threshold is the 50% point between 0 and the maximum firing rate. (**d**) Speaker layout reference.

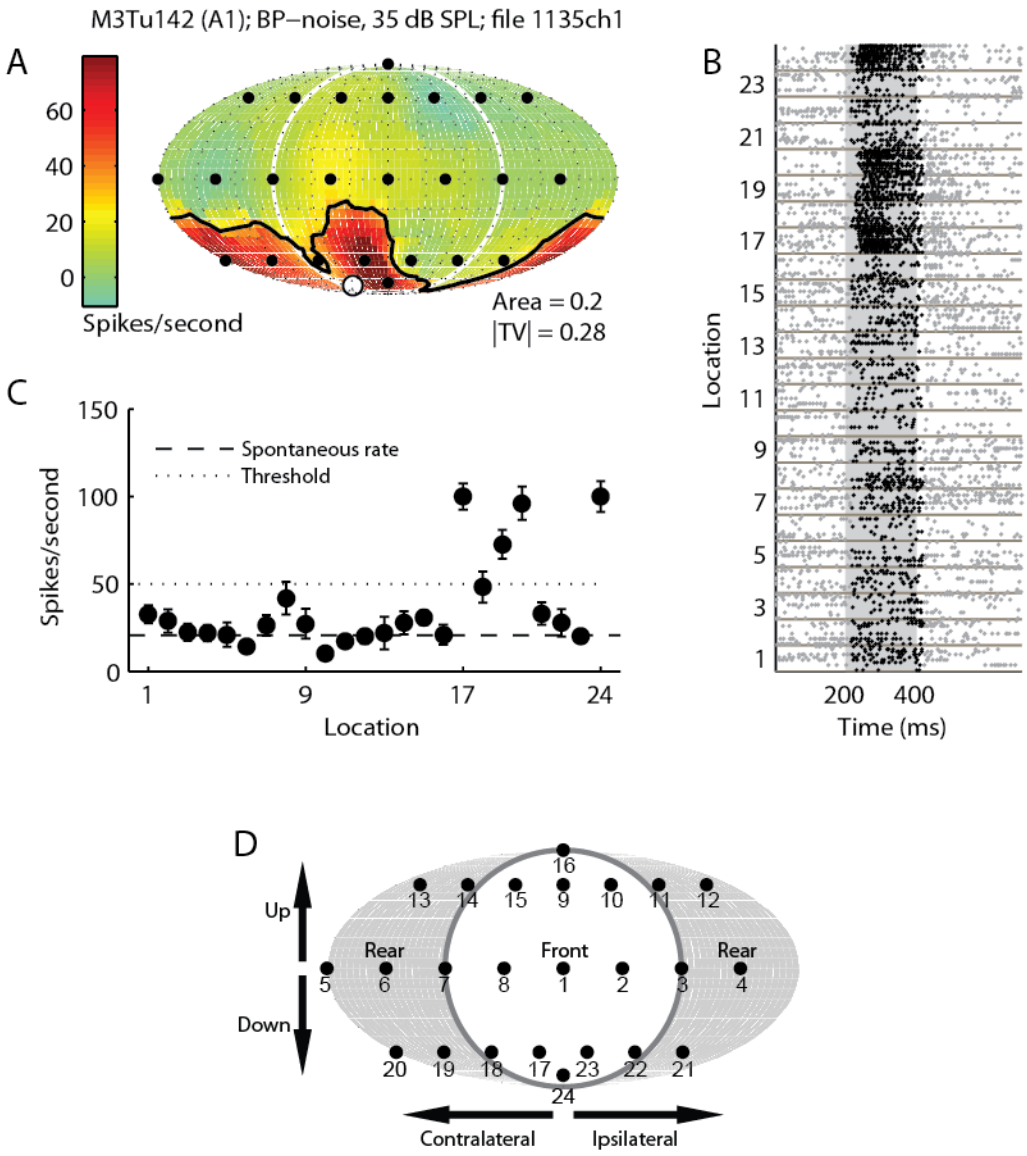


Figure 3.3.14. Spatial receptive field examples (14). Example unit tuned below the horizon. (**a**) Relative firing rates displayed on a Fournier projection of space. (**b**) Peristimulus time histogram (PSTH) raster plot. Analysis windows are 15 milliseconds post-stimulus onset to 20 milliseconds post-stimulus offset. (**c**) Firing rates at different speaker locations. Error bars represent standard error of the mean. The threshold is the 50% point between 0 and the maximum firing rate. (**d**) Speaker layout reference.

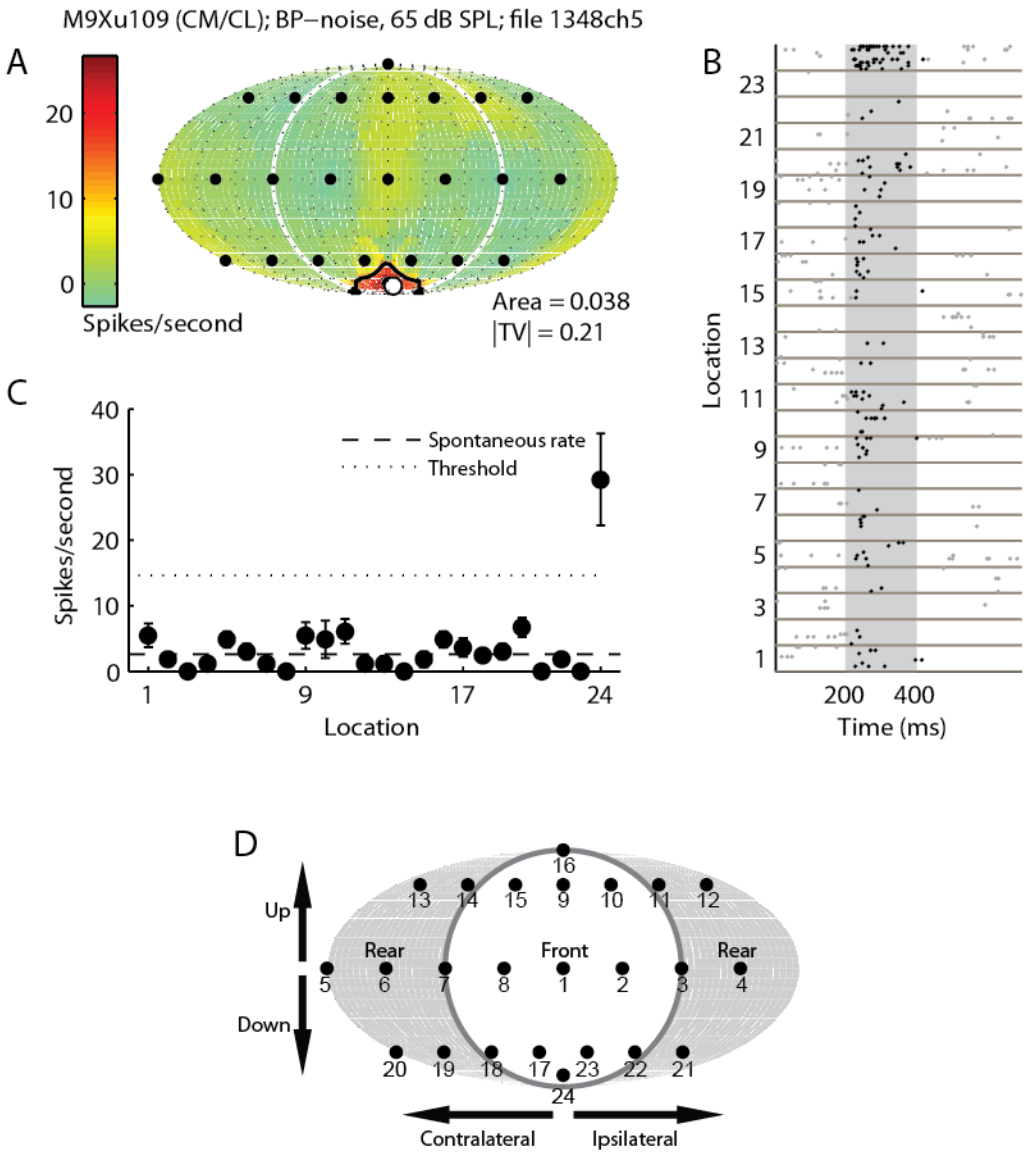


Figure 3.3.15. Spatial receptive field examples (15). Example unit (2) tuned below the horizon. (**a**) Relative firing rates displayed on a Fournier projection of space. (**b**) Peristimulus time histogram (PSTH) raster plot. Analysis windows are 15 milliseconds post-stimulus onset to 20 milliseconds post-stimulus offset. (**c**) Firing rates at different speaker locations. Error bars represent standard error of the mean. The threshold is the 50% point between 0 and the maximum firing rate. (**d**) Speaker layout reference.

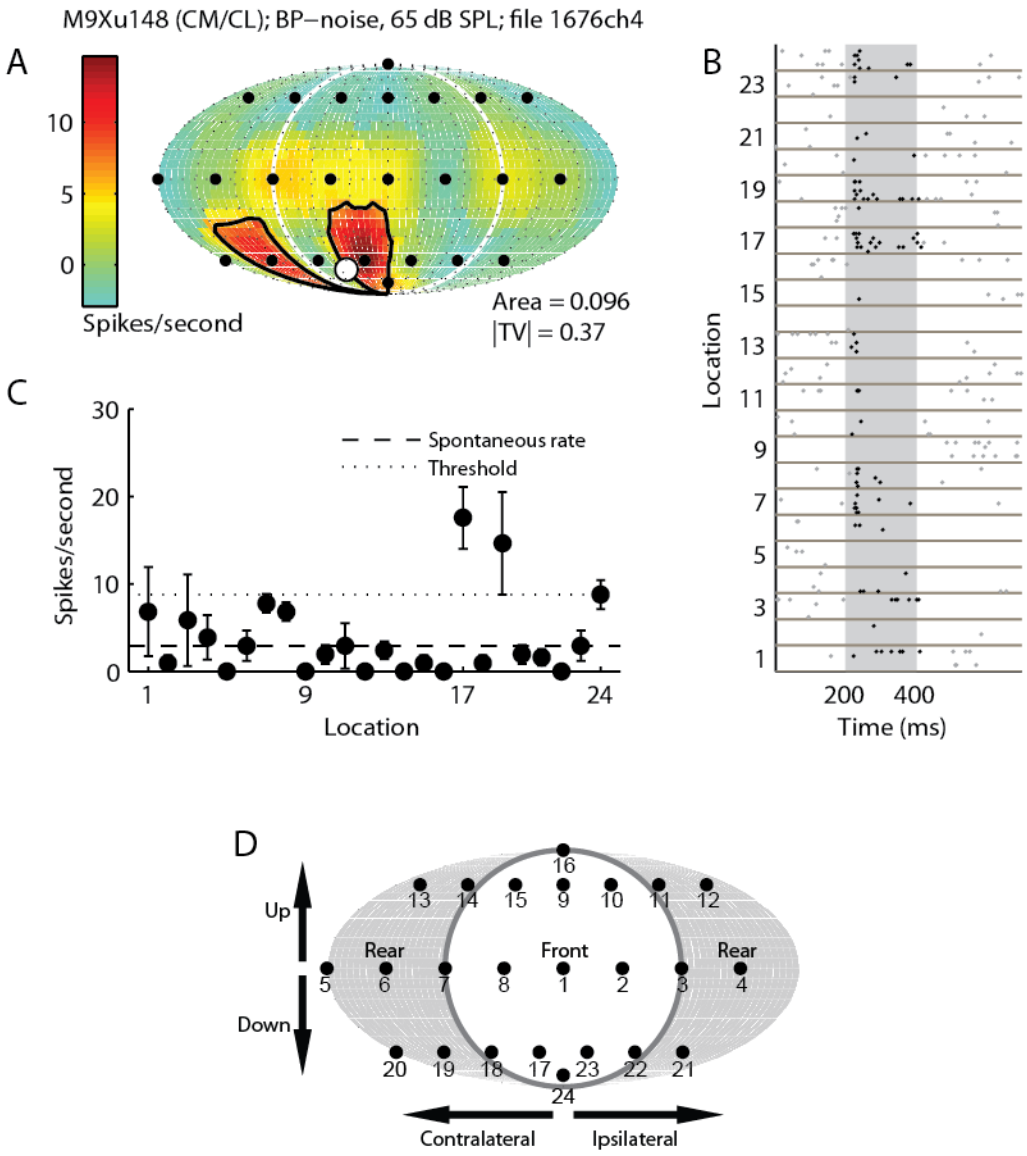


Figure 3.3.16. Spatial receptive field examples (16). Example unit (3) tuned below the horizon. (**a**) Relative firing rates displayed on a Fournier projection of space. (**b**) Peristimulus time histogram (PSTH) raster plot. Analysis windows are 15 milliseconds post-stimulus onset to 20 milliseconds post-stimulus offset. (**c**) Firing rates at different speaker locations. Error bars represent standard error of the mean. The threshold is the 50% point between 0 and the maximum firing rate. (**d**) Speaker layout reference.

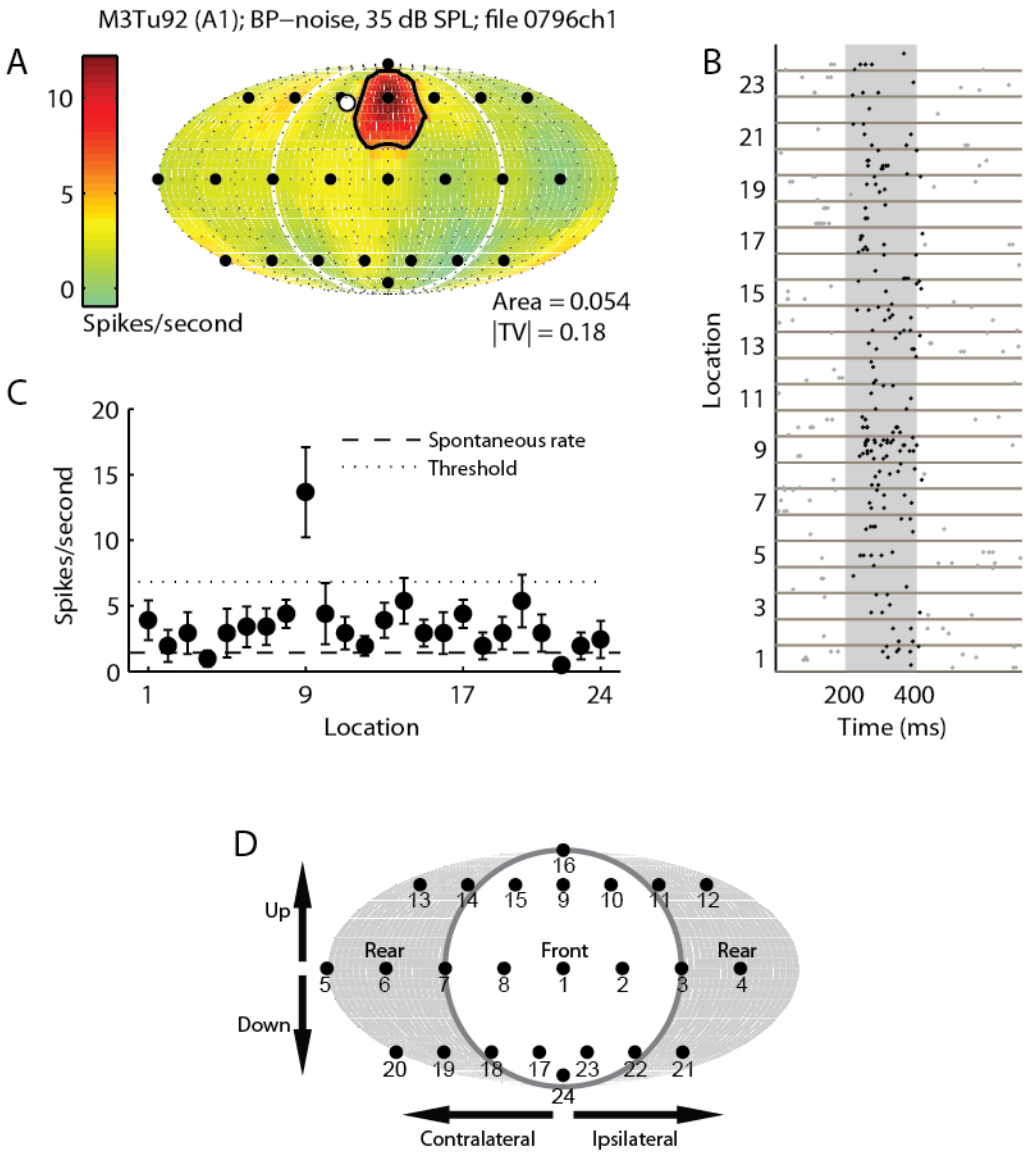


Figure 3.3.17. Spatial receptive field examples (17). Example unit tuned above the horizon. (**a**) Relative firing rates displayed on a Fournier projection of space. (**b**) Peristimulus time histogram (PSTH) raster plot. Analysis windows are 15 milliseconds post-stimulus onset to 20 milliseconds post-stimulus offset. (**c**) Firing rates at different speaker locations. Error bars represent standard error of the mean. The threshold is the 50% point between 0 and the maximum firing rate. (**d**) Speaker layout reference.

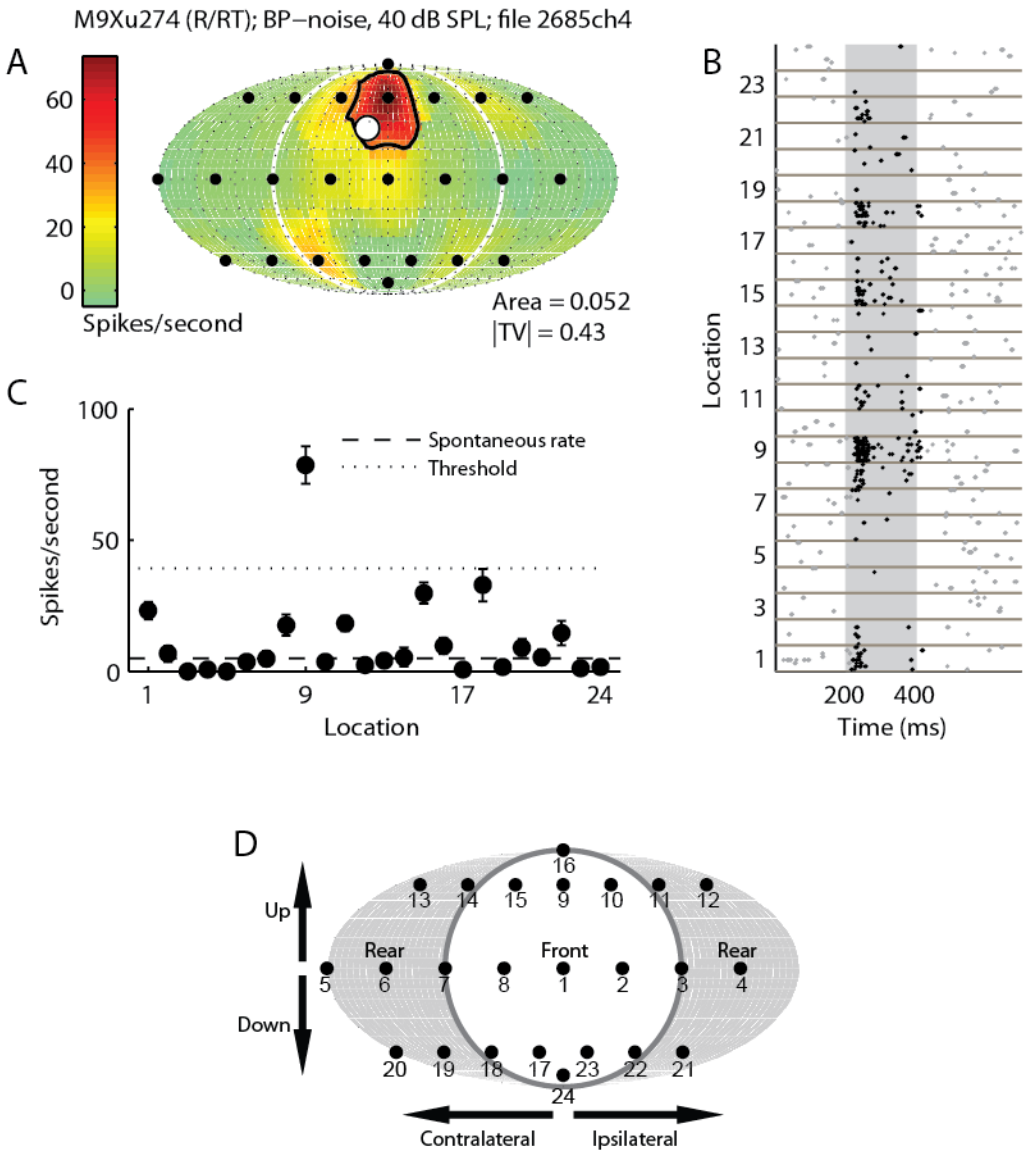


Figure 3.3.18. Spatial receptive field examples (18). Example unit (2) tuned above the horizon. (**a**) Relative firing rates displayed on a Fournier projection of space. (**b**) Peristimulus time histogram (PSTH) raster plot. Analysis windows are 15 milliseconds post-stimulus onset to 20 milliseconds post-stimulus offset. (**c**) Firing rates at different speaker locations. Error bars represent standard error of the mean. The threshold is the 50% point between 0 and the maximum firing rate. (**d**) Speaker layout reference.

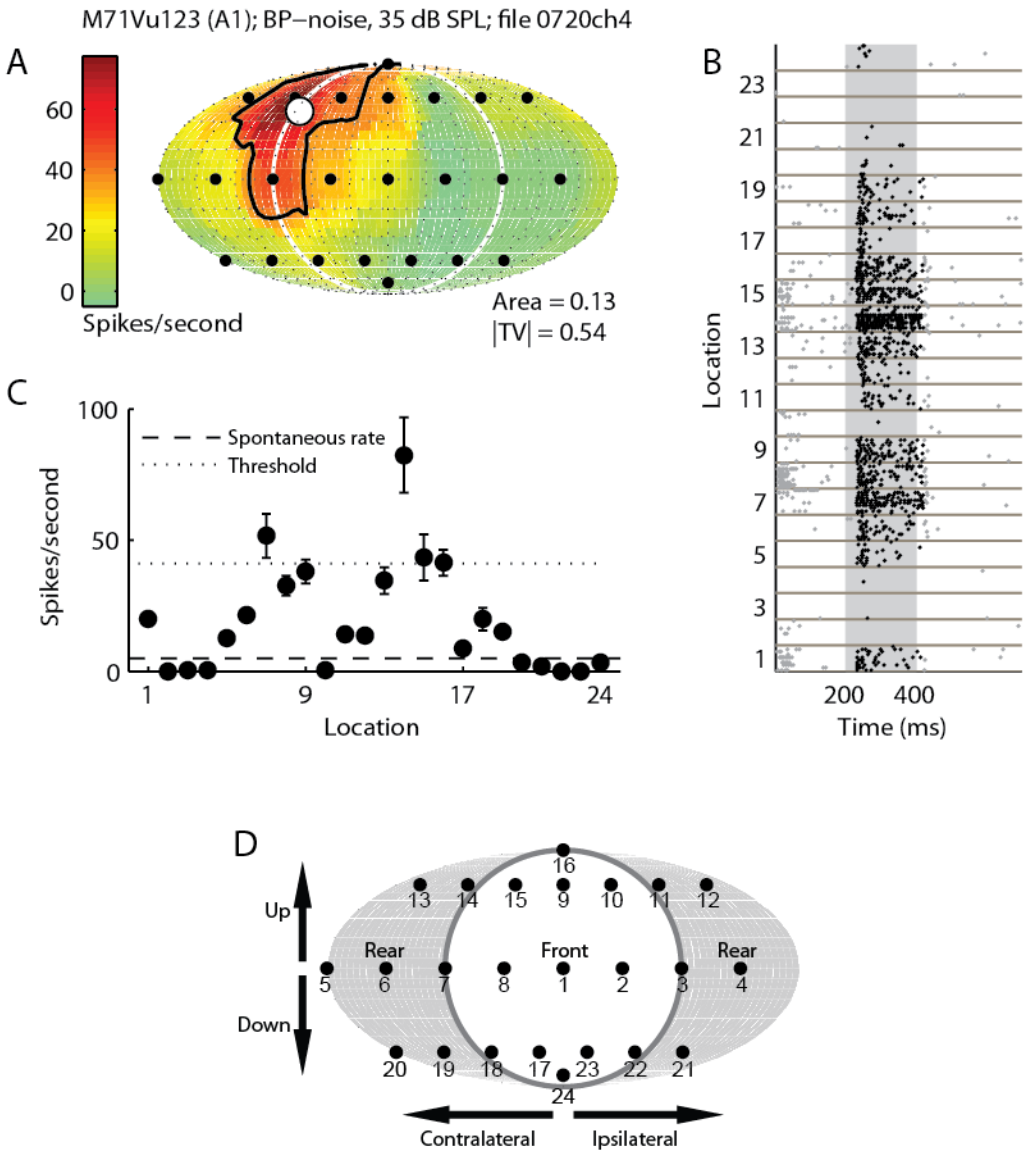


Figure 3.3.19. Spatial receptive field examples (19). Example unit (3) tuned above the horizon. (**a**) Relative firing rates displayed on a Fournier projection of space. (**b**) Peristimulus time histogram (PSTH) raster plot. Analysis windows are 15 milliseconds post-stimulus onset to 20 milliseconds post-stimulus offset. (**c**) Firing rates at different speaker locations. Error bars represent standard error of the mean. The threshold is the 50% point between 0 and the maximum firing rate. (**d**) Speaker layout reference.

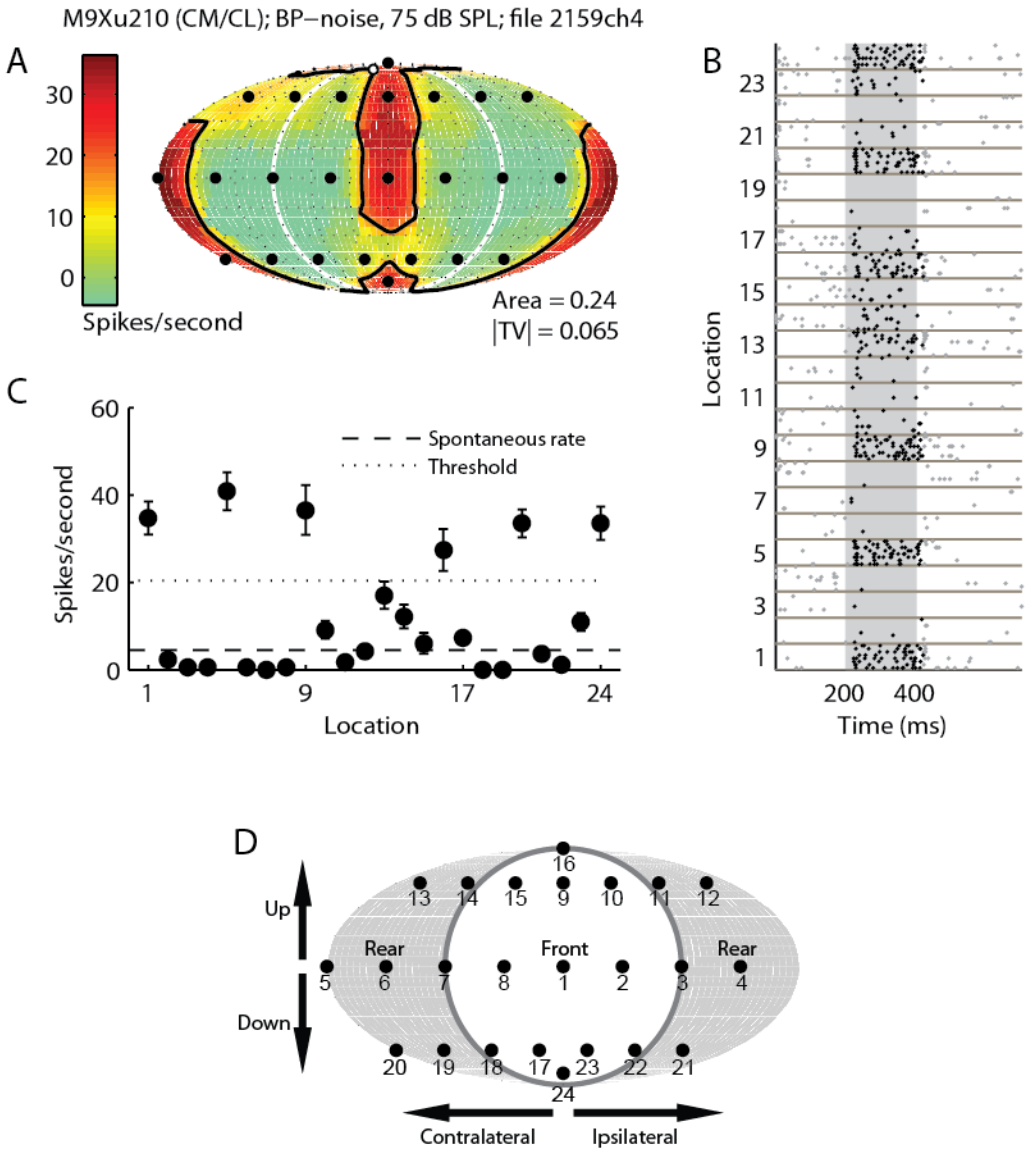


Figure 3.3.20. Spatial receptive field examples (20). Example unit tuned to the midline. (**a**) Relative firing rates displayed on a Fournier projection of space. (**b**) Peristimulus time histogram (PSTH) raster plot. Analysis windows are 15 milliseconds post-stimulus onset to 20 milliseconds post-stimulus offset. (**c**) Firing rates at different speaker locations. Error bars represent standard error of the mean. The threshold is the 50% point between 0 and the maximum firing rate. (**d**) Speaker layout reference.

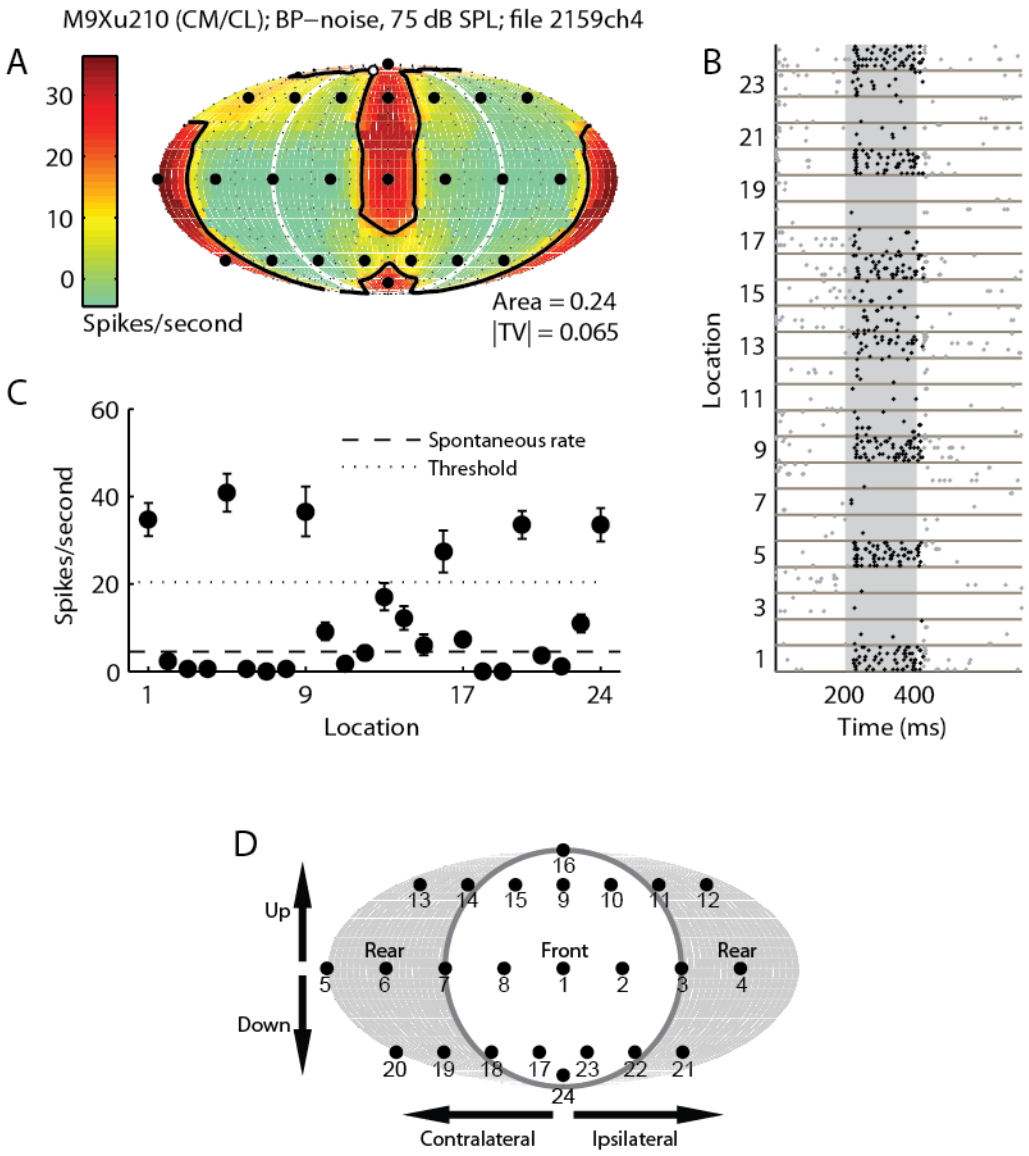


Figure 3.3.21. Spatial receptive field examples (21). Example unit (2) tuned to the midline. (**a**) Relative firing rates displayed on a Fournier projection of space. (**b**) Peristimulus time histogram (PSTH) raster plot. Analysis windows are 15 milliseconds post-stimulus onset to 20 milliseconds post-stimulus offset. (**c**) Firing rates at different speaker locations. Error bars represent standard error of the mean. The threshold is the 50% point between 0 and the maximum firing rate. (**d**) Speaker layout reference.

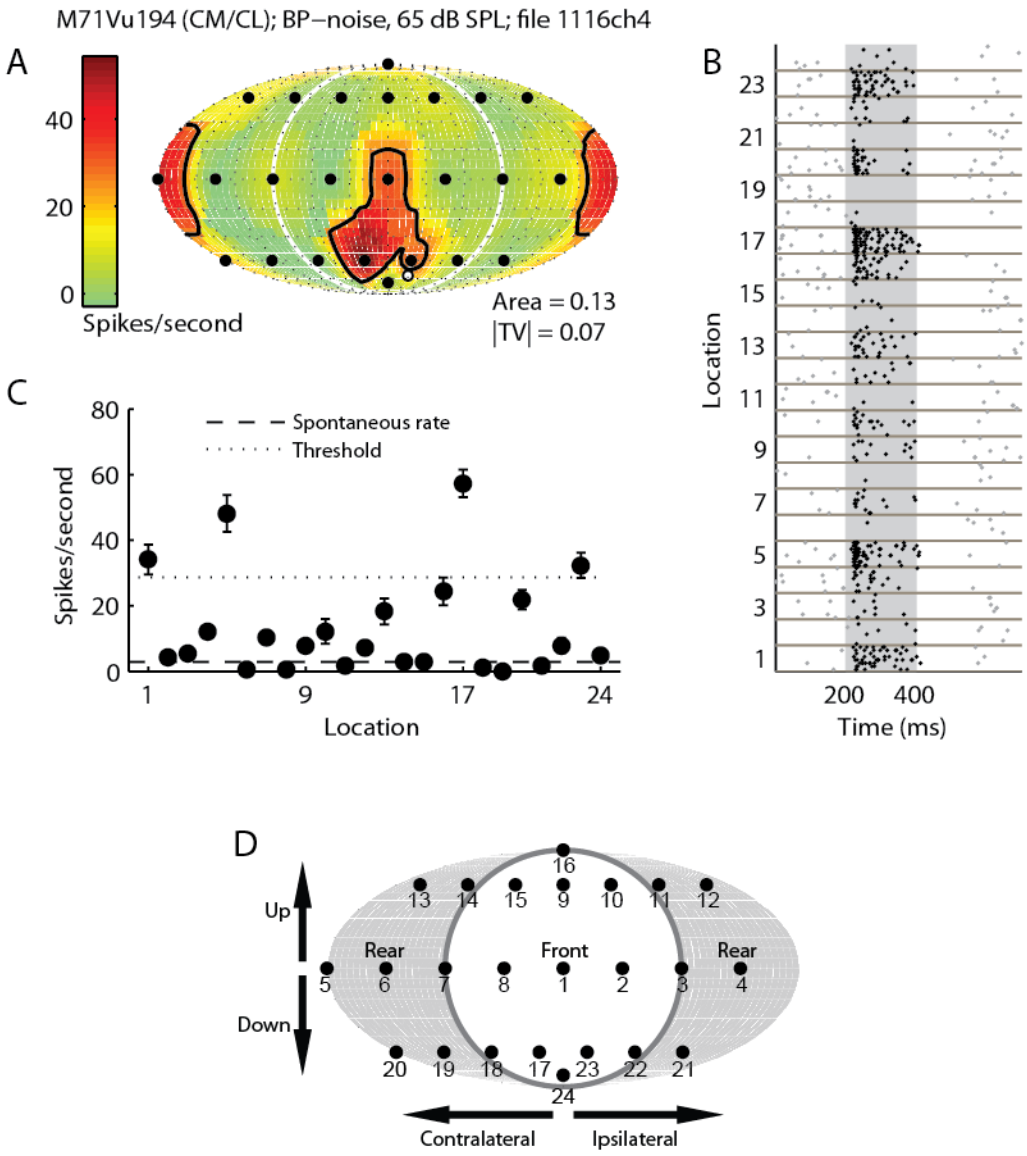
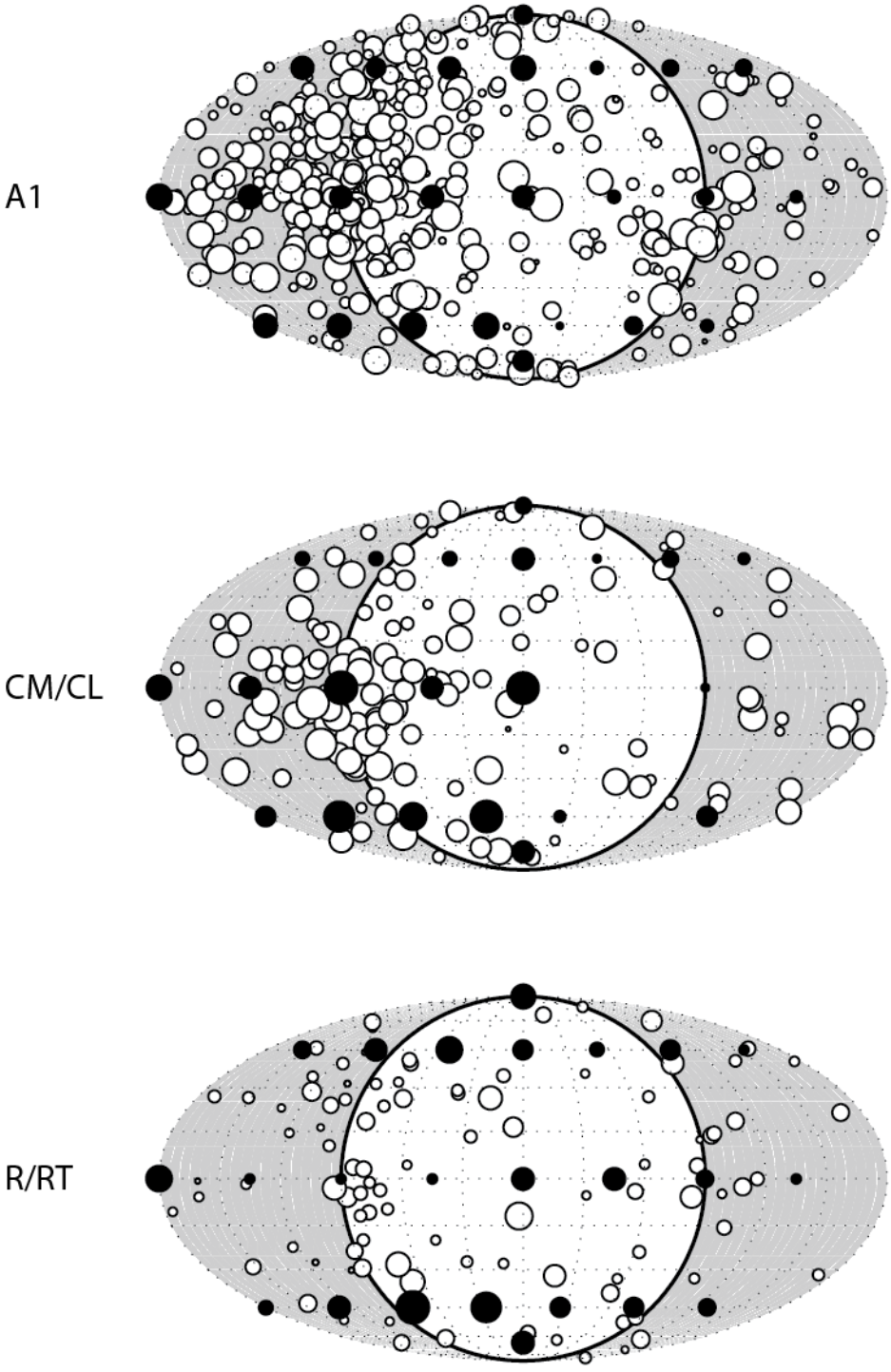
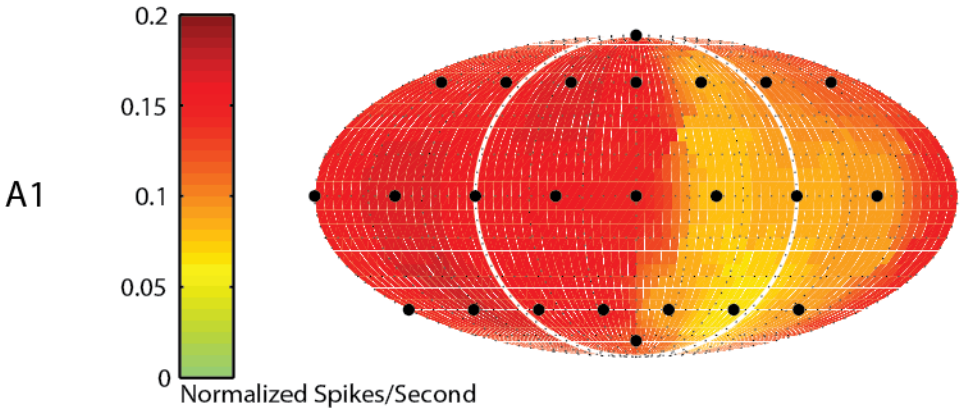
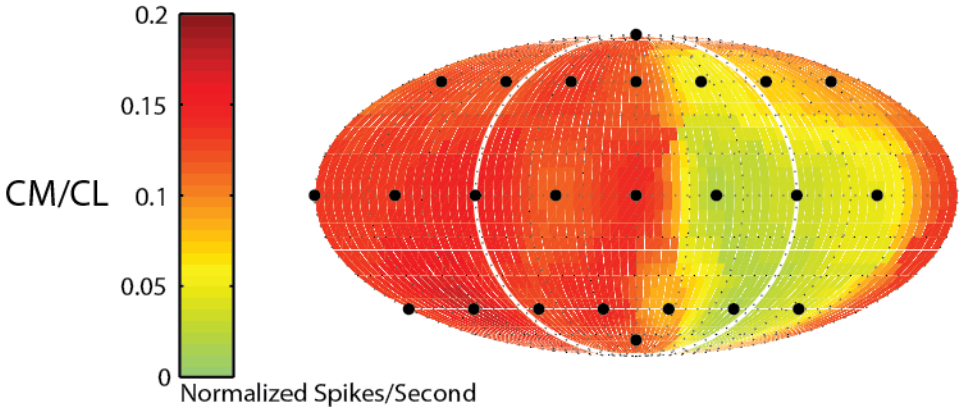


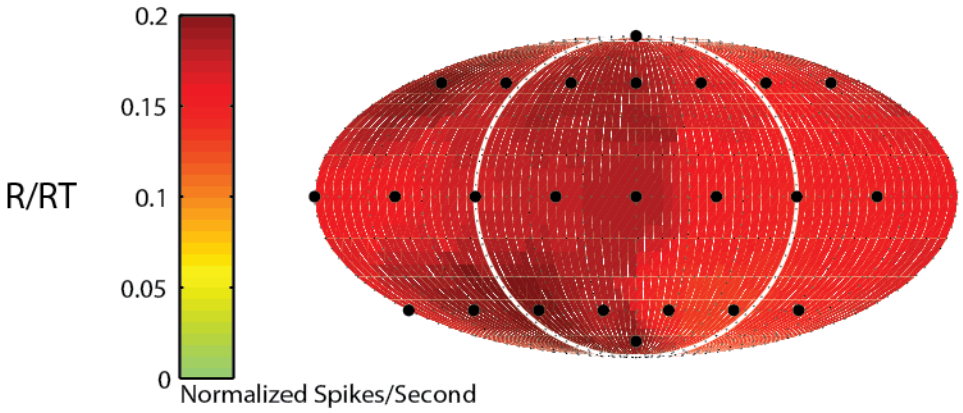
Figure 3.3.22. Spatial receptive field examples (22). Example unit (3) tuned to the midline. (**a**) Relative firing rates displayed on a Fournier projection of space. (**b**) Peristimulus time histogram (PSTH) raster plot. Analysis windows are 15 milliseconds post-stimulus onset to 20 milliseconds post-stimulus offset. (**c**) Firing rates at different speaker locations. Error bars represent standard error of the mean. The threshold is the 50% point between 0 and the maximum firing rate. (**d**) Speaker layout reference.

Figure 3.4.1-2. Distribution of spatial tuning in auditory cortex areas A1, CM/CL, and R/RT. (**1**) Histogram of best locations and distributions of tuning vector locations. The number of neurons preferring each individual location is indicated by the size of filled black circles located at speaker locations (see figure 3.1). White circles indicate tuning vector locations; their magnitudes (a larger tuning vector magnitude indicates a more directional receptive field) are indicated by the size of the circles. All areas have a broad distribution of tuning vectors and best locations located in all regions of space. (**b**) Mean of normalized firing rates across each area. Rates were higher in the contralateral hemifield than the ipsilateral hemifield, but this trend was weaker in A1 and R/RT than in CM/CL. Rates within the contralateral hemifield did were not well correlated with lateral angle (A1: $r = 0.04$, $p < 0.002$; CM/CL: $r = -0.02$, $p = 0.3$; R/RT: $r = 0.008$, $p = 0.8$)


A1
CM/CL
R/RT


A1
0.2
0.15
0.1
0.05
0
Normalized Spikes/Second


CM/CL
0.2
0.15
0.1
0.05
0
Normalized Spikes/Second


R/RT
0.2
0.15
0.1
0.05
0
Normalized Spikes/Second

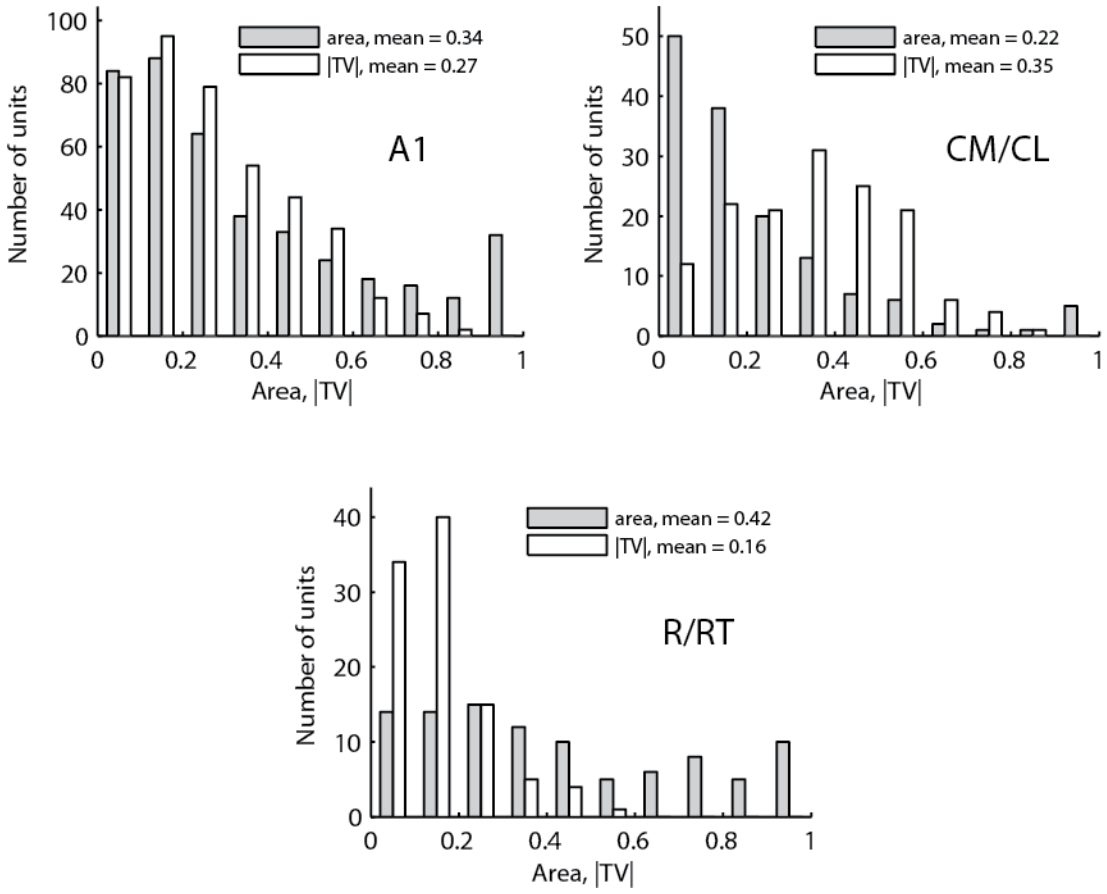


Figure 3.5. Distribution of tuning area and vector magnitude in A1, CM/CL, and R/RT. Tuning area was highly variable, and skewed towards relatively low values in A1 and particularly in CM/CL. This effect of cortical area on tuning area was highly significant (ANOVA: $p << 0.001$), with R/RT having the largest tuning areas (mean = 0.43), followed by A1 (mean = 0.34), and CM/CL (mean = 0.22). Tuning vector magnitude is sensitive to shape and dispersion as well as extent of spatial receptive fields (see figure 3e). Tuning vector magnitude followed the same trend vs. area as tuning width (ANOVA: $p << 0.001$), with CM/CL having the most directional neurons (mean = 0.35), followed by A1 (mean = 0.27), and R/RT (0.16).

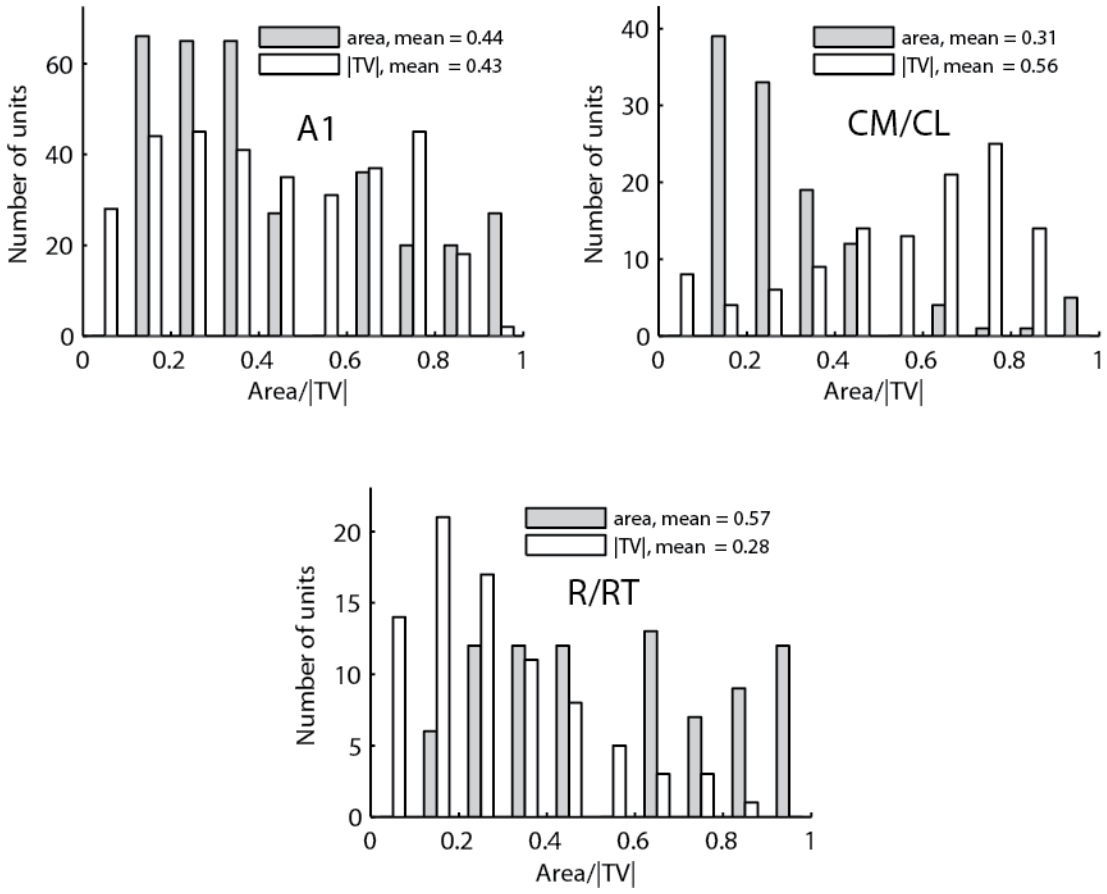


Figure 3.6. Distribution of tuning area and vector magnitude: azimuth only. In all areas, tuning areas were significantly larger when only azimuth was considered compared with those when the full spatial field was considered. Tuning vectors were larger, but this is a side effect of the reduced dimensionality on the vector calculation.

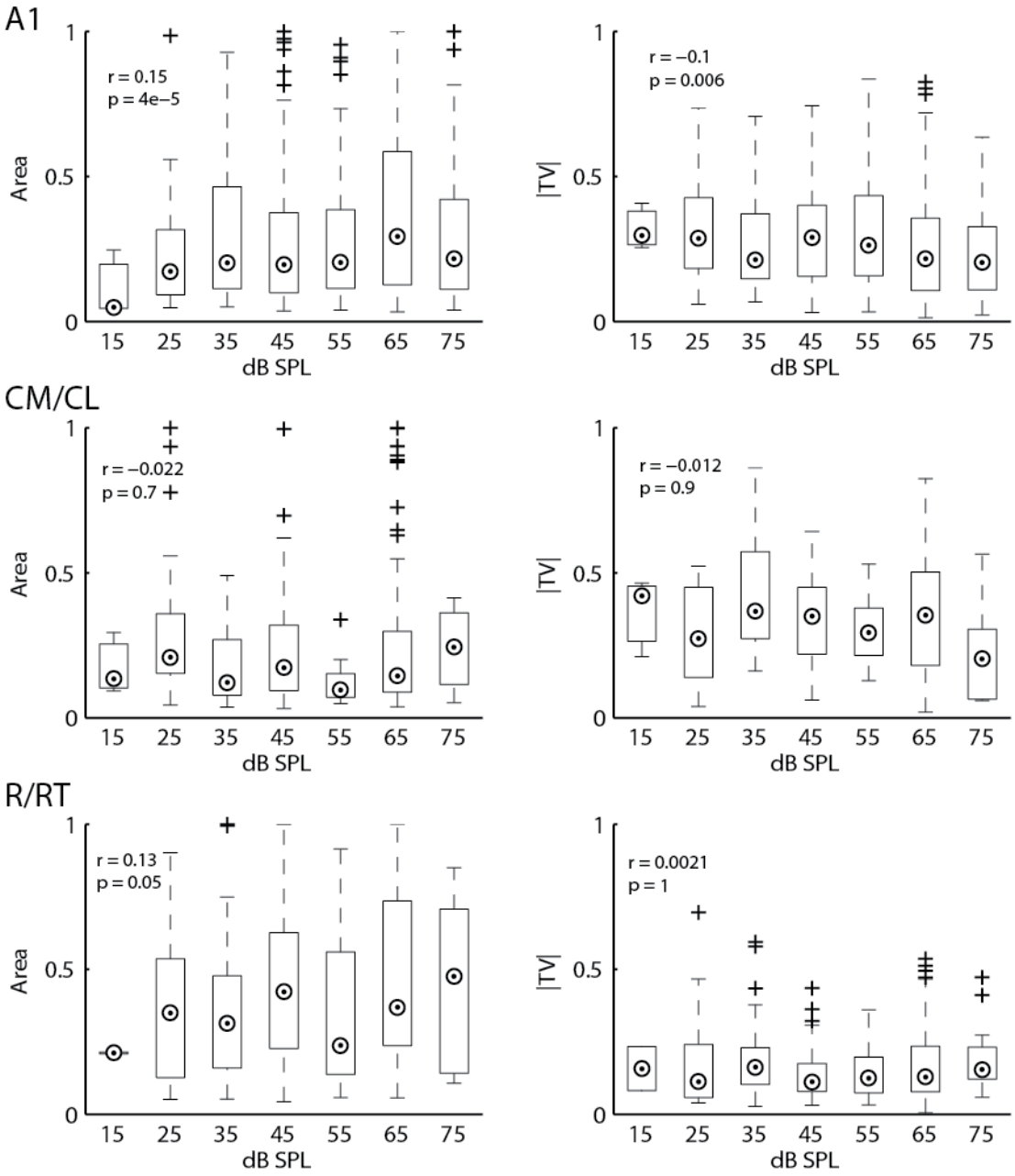


Figure 3.7. Population spatial selectivity vs. sound pressure level. Analysis of the effects of increasing sound level on the distribution of tuning selectivity measures. In A1 and R/RT, increasing sound level decreased spatial selectivity, but CM/CL did not show the same effects.

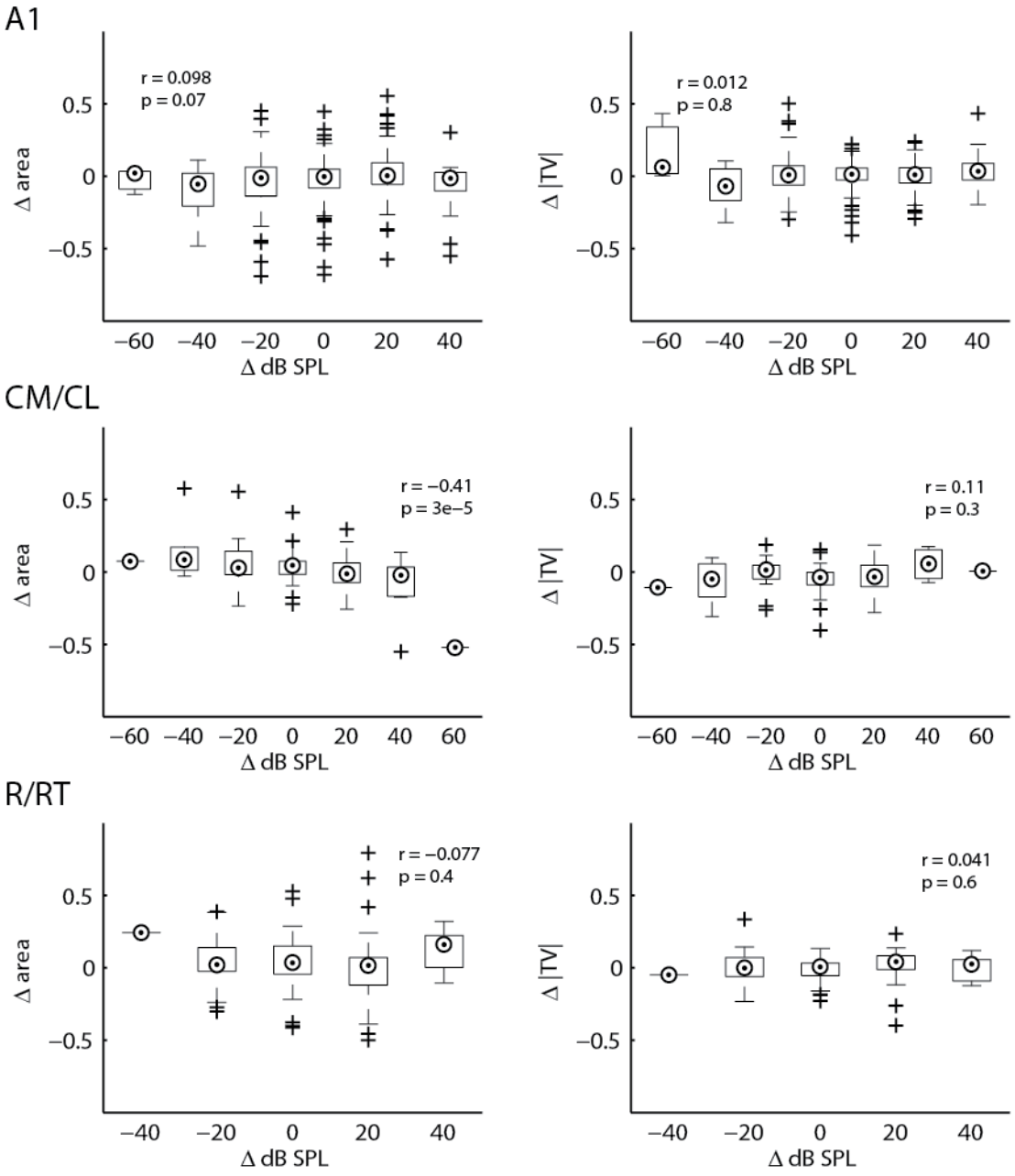


Figure 3.8. Within unit spatial selectivity vs. sound pressure level. Within neuron analysis of level effects on tuning area and vector magnitude, referenced to best SRF level. There were no correlations between ΔdB SPL and tuning area shift in areas A1 or R/RT, but a significant negative correlation was observed between tuning area shift and sound level shift in CM/CL units. Data points at ΔdB SPL = 0 represent instances when the same level was tested multiple times in the same neuron.

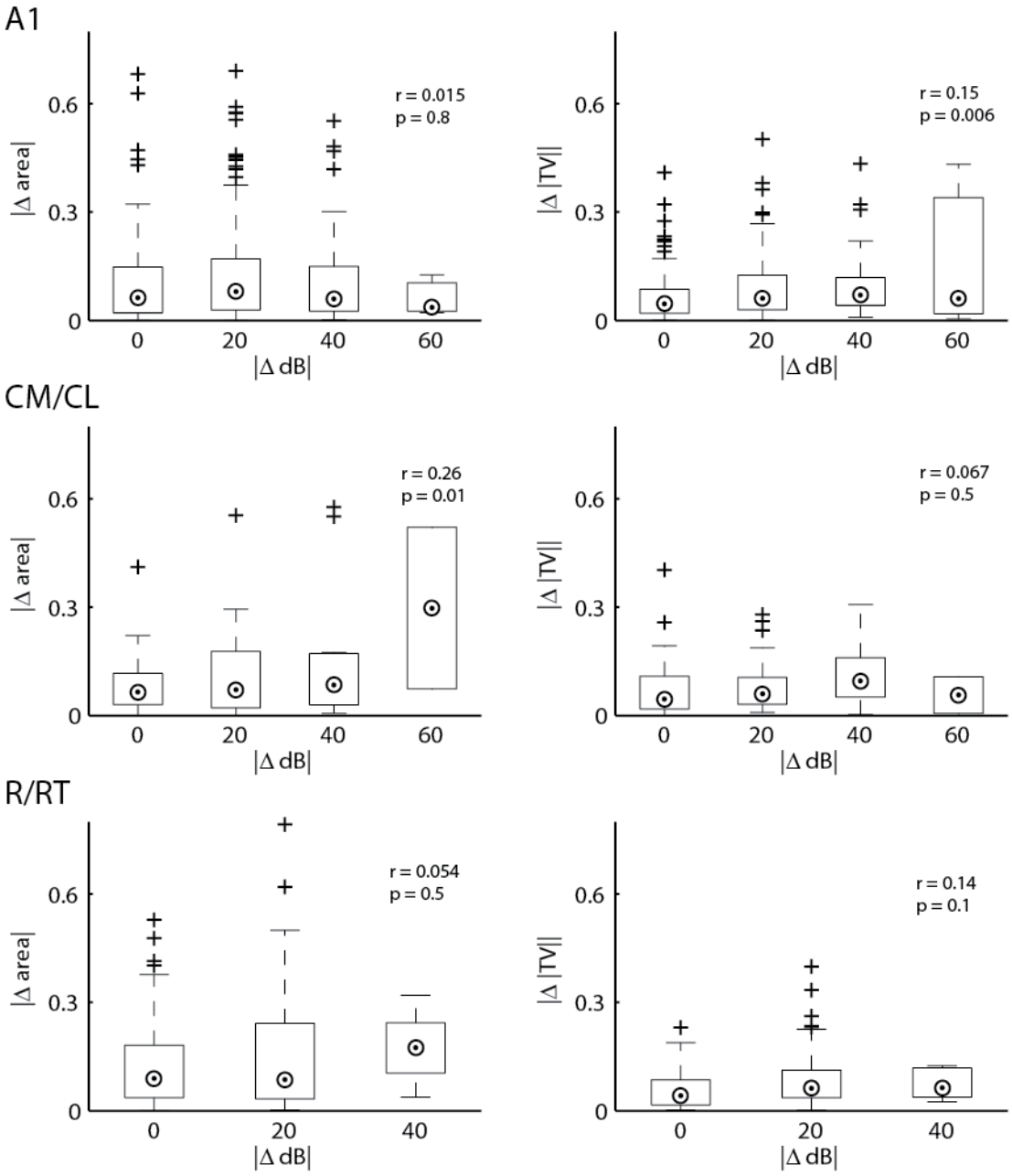


Figure 3.9. Spatial selectivity stability across pressure level. Effect of the absolute value of level changes on the absolute value of tuning width change, both referenced to the best SRF. There was no significant effect in R/RT and a weak effect in A1 for tuning vector magnitude and in CM/CL for tuning area. The relationship in CM/CL is expected from the negative correlation between sound level and tuning area.

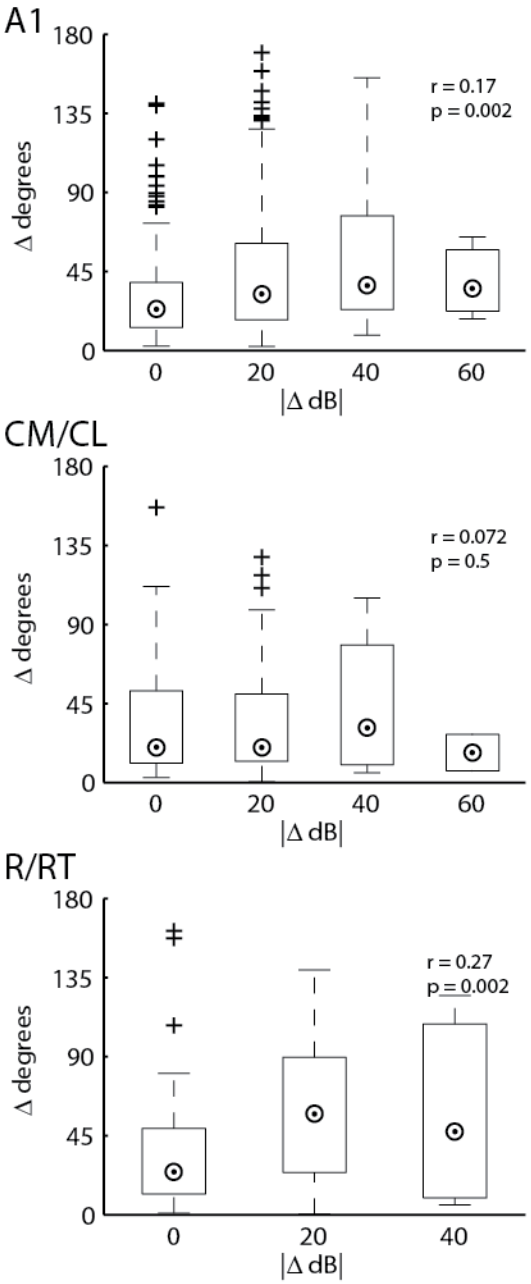


Figure 3.10. Tuning vector stability vs. sound pressure level. Change in tuning vector (in degrees) as a function of ΔdB SPL. Weak effects of ΔdB SPL on tuning vector shift magnitude were observed in A1 and R/RT but not CM/CL. The smallest tuning vector shift magnitudes were observed in CM/CL (ANCOVA, $p < .05$).

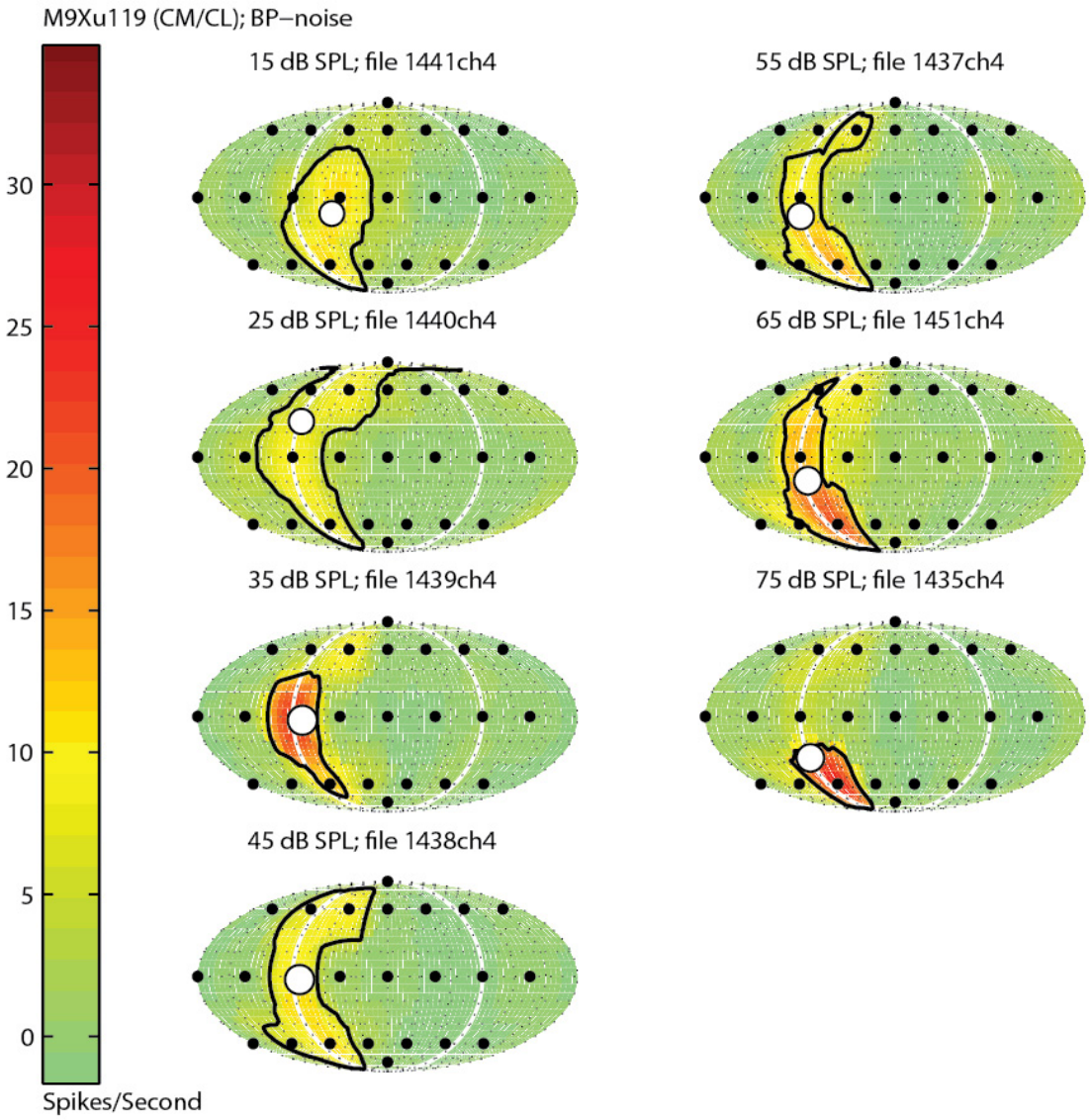


Figure 3.11.1. Examples of spatial tuning stability across sound level. Many neurons showed spatial tuning which was highly conserved across sound level.

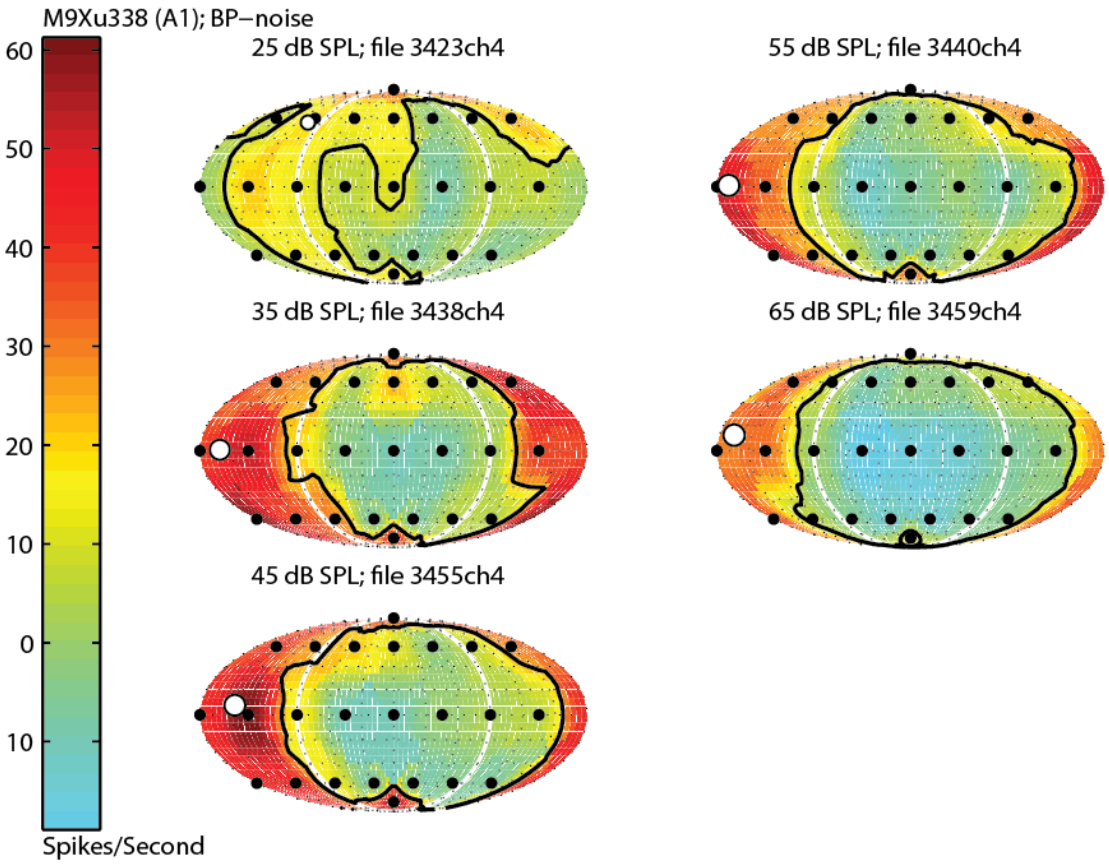


Figure 3.11.2. Examples of spatial tuning stability across sound level. Many neurons showed spatial tuning which was highly conserved across sound level (example 2).

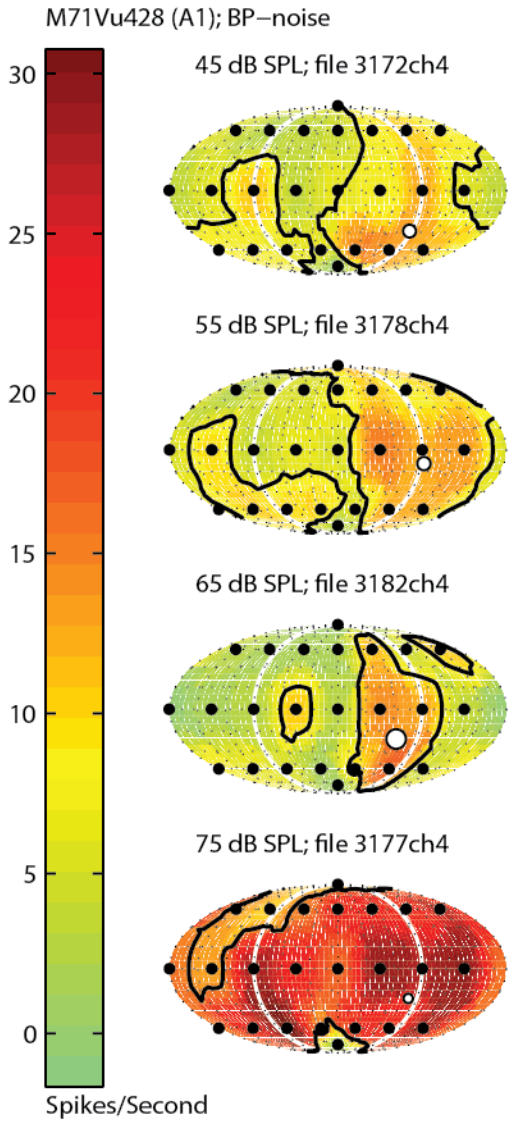


Figure 3.11.3. Examples of spatial tuning stability across sound level. A minority of neurons displayed substantially expanded spatial tuning at high sound levels.

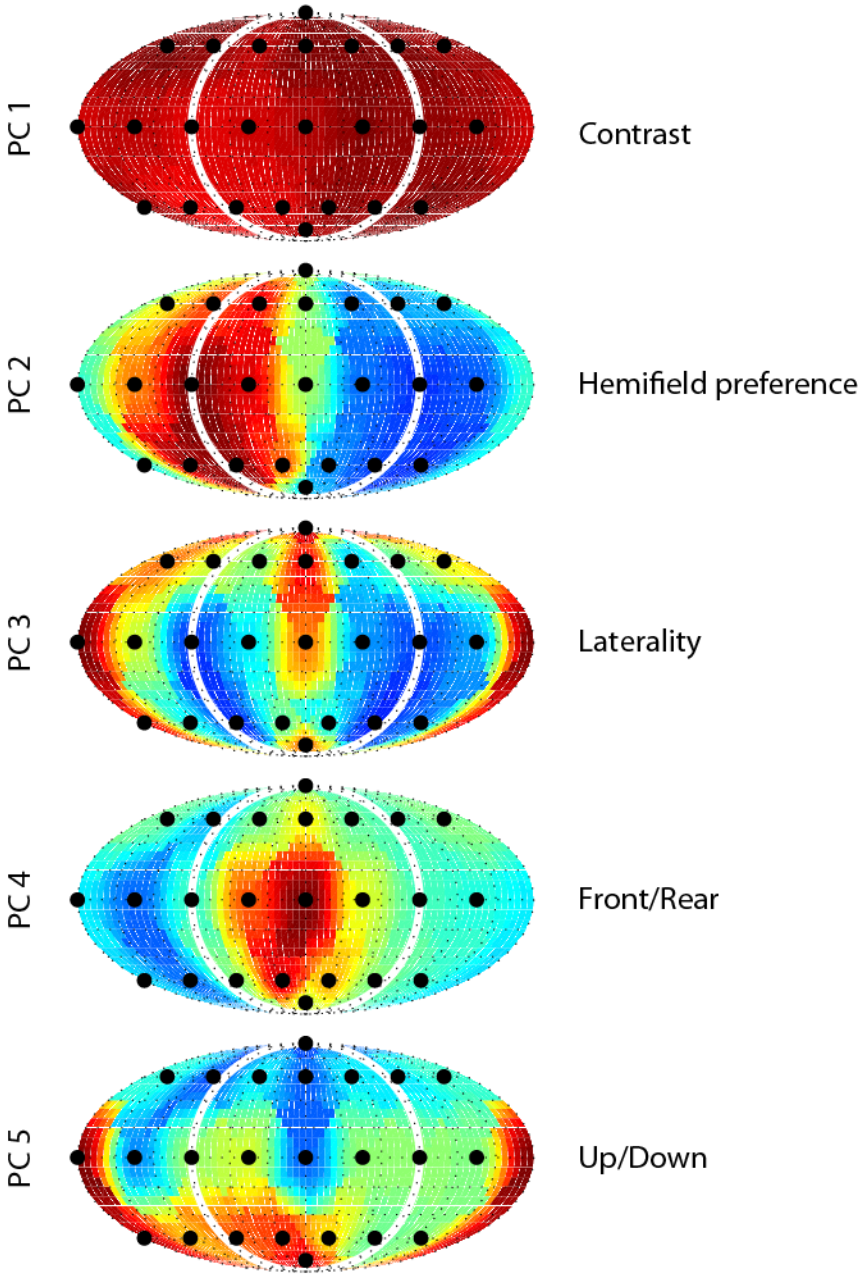


Figure 3.12. SRF principal component analysis. Spatial receptive fields varied along biologically relevant dimensions. The first principal component, accounting for 40% of receptive field variance, approximates the non-directional component of neural responses, thus is referred to as "contrast." The second and third components, reflecting a neurons' hemifield and laterality preference, together describe azimuth tuning. The fourth and fifth components represent front/back and up/down variation. Variation in azimuth makes up a larger fraction of variance than median plane tuning (18 % vs. 10%).

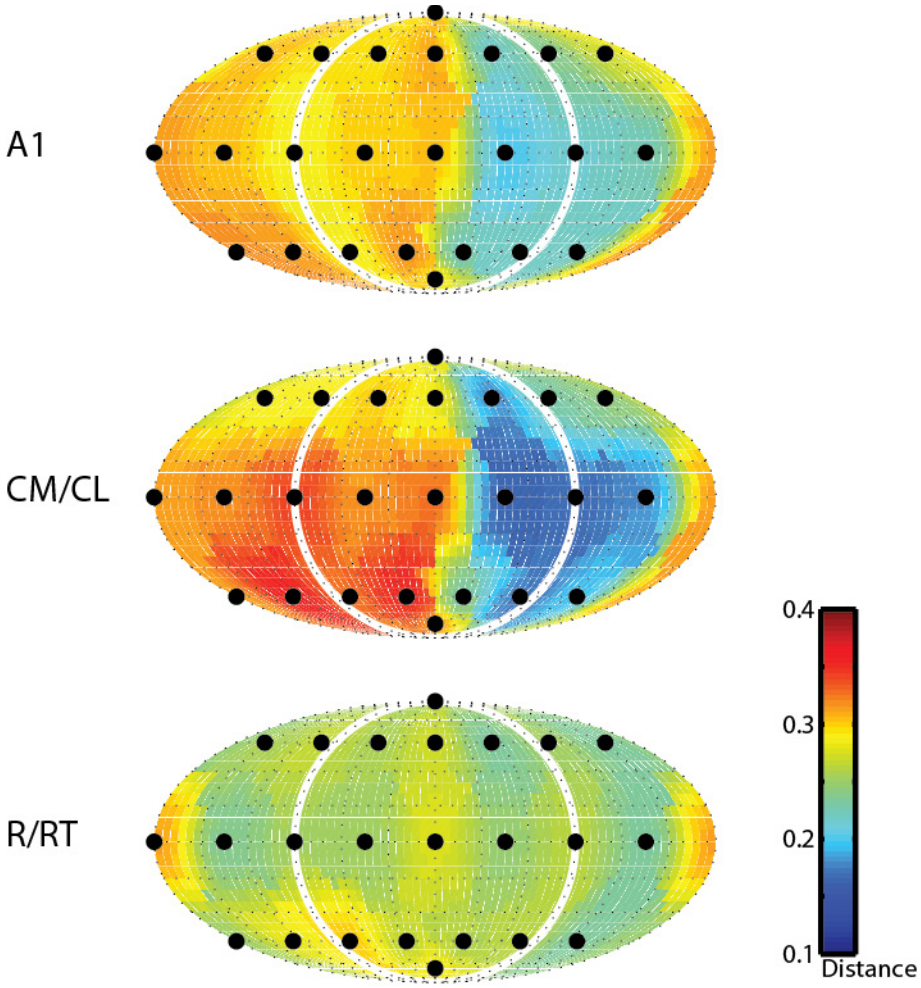


Figure 3.13. Mean Euclidean distance. Euclidean distance, normalized by population size, to the nearest 4 locations (measured by firing rate) for each location. Euclidean distance between the nearest locations increased in the contralateral hemifield and decreased in the ipsilateral hemifield from R/RT to A1to CM/CL (repeated measures ANOVA, $p << .001$). In the median plane, there was a significant effect of area (repeated measures ANOVA, $p << .001$), but there was no significant difference between A1 and CM/CL (repeated measures ANOVA, $p = .97$).

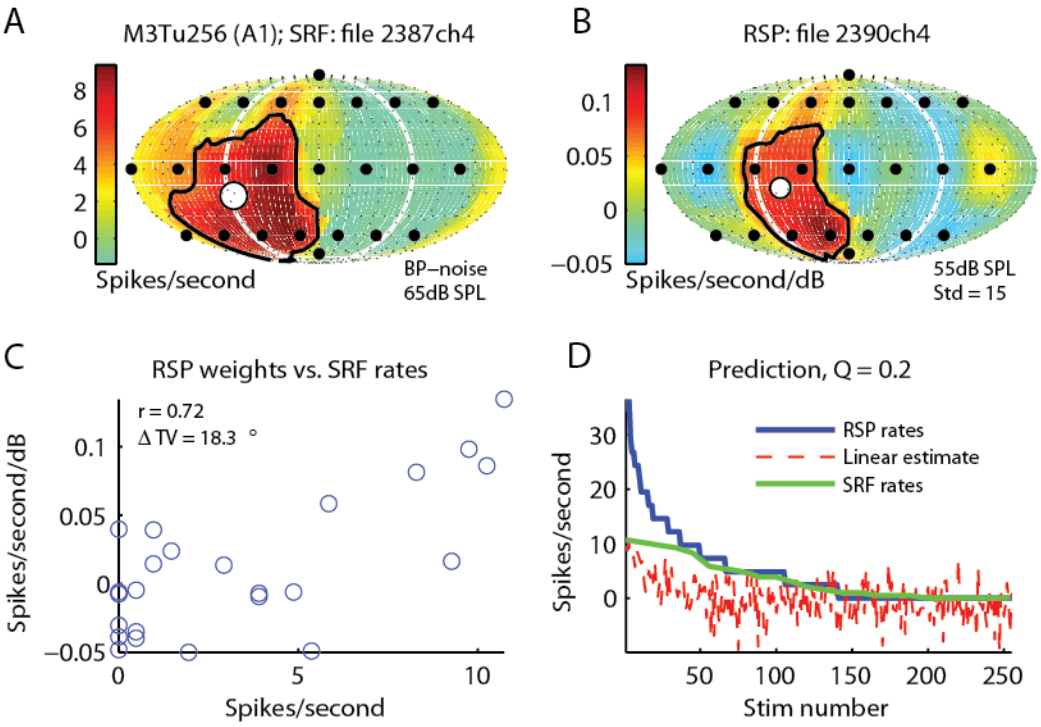


Figure 3.14.1. Comparison of 1- and 24-speaker receptive fields. An example neuron with good agreement between the two conditions. (**a**) Spatial receptive field measured in the 1-speaker condition. (**b**) Spatial weighting function calculated using responses to random spatial profile (RSP) stimuli. (**c**) RSP weights vs. SRF weights. (**d**) Responses to RSP stimuli ordered from highest to lowest (blue), the linear prediction using the weighting function and the RSP level matrix (red), and the 1-speaker responses (green), ordered from highest to lowest.

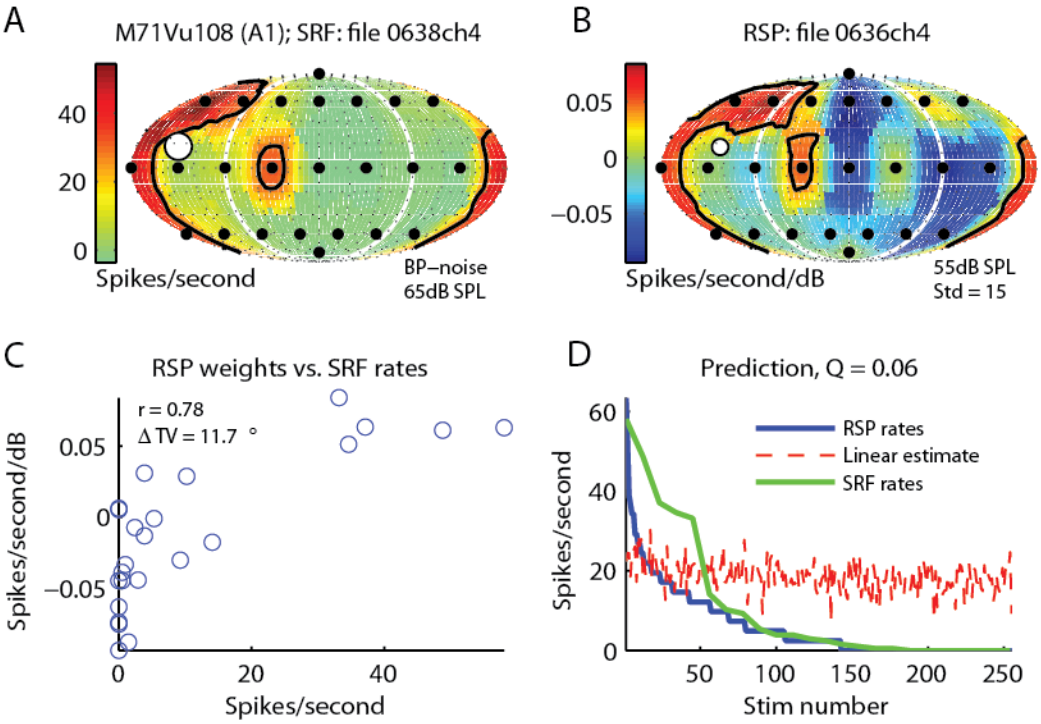


Figure 3.14.2. Comparison of 1- and 24-speaker receptive fields (2). An example neuron (2) with good agreement between the two conditions. (**a**) Spatial receptive field measured in the 1-speaker condition. (**b**) Spatial weighting function calculated using responses to random spatial profile (RSP) stimuli. (**c**) RSP weights vs. SRF weights. (**d**) Responses to RSP stimuli ordered from highest to lowest (blue), the linear prediction using the weighting function and the RSP level matrix (red), and the 1-speaker responses (green), ordered from highest to lowest.

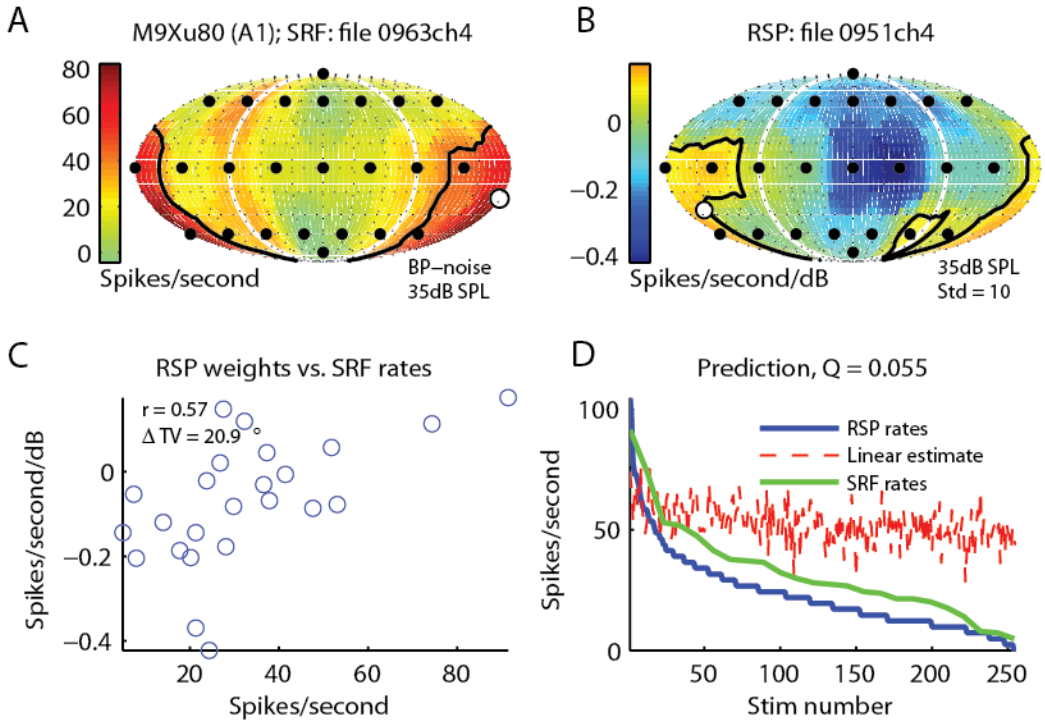


Figure 3.14.3. Comparison of 1- and 24-speaker receptive fields (3). An example neuron (3) with good agreement between the two conditions. (**a**) Spatial receptive field measured in the 1-speaker condition. (**b**) Spatial weighting function calculated using responses to random spatial profile (RSP) stimuli. (**c**) RSP weights vs. SRF weights. (**d**) Responses to RSP stimuli ordered from highest to lowest (blue), the linear prediction using the weighting function and the RSP level matrix (red), and the 1-speaker responses (green), ordered from highest to lowest.

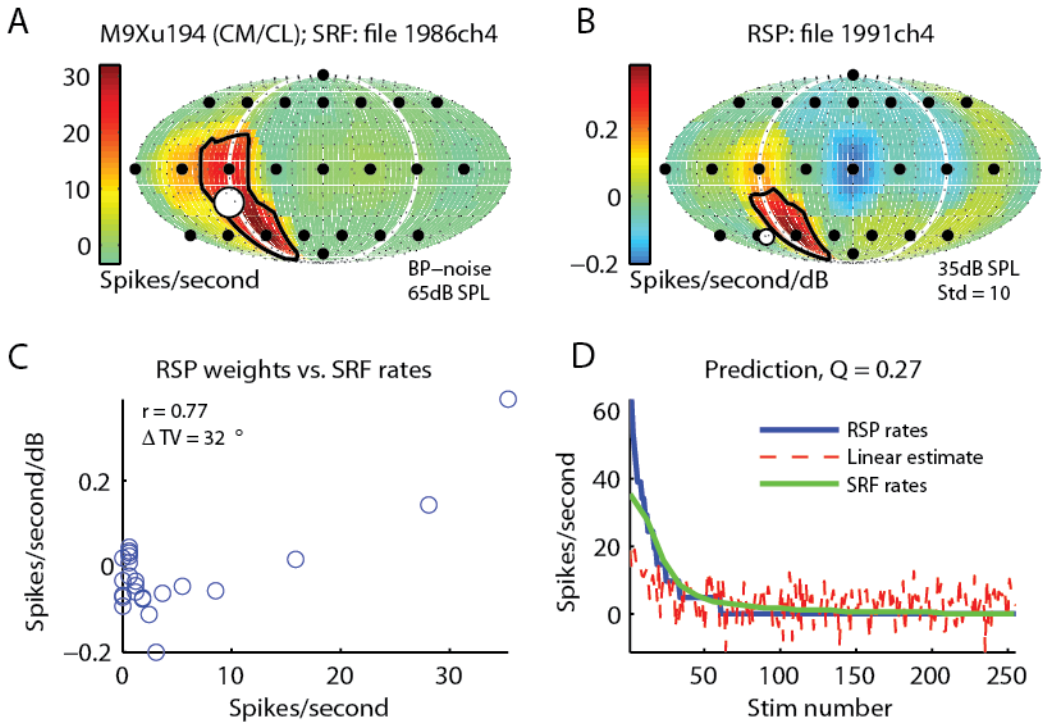


Figure 3.14.4. Comparison of 1- and 24-speaker receptive fields (4). An example neuron (4) with good agreement between the two conditions. (**a**) Spatial receptive field measured in the 1-speaker condition. (**b**) Spatial weighting function calculated using responses to random spatial profile (RSP) stimuli. (**c**) RSP weights vs. SRF weights. (**d**) Responses to RSP stimuli ordered from highest to lowest (blue), the linear prediction using the weighting function and the RSP level matrix (red), and the 1-speaker responses (green), ordered from highest to lowest.

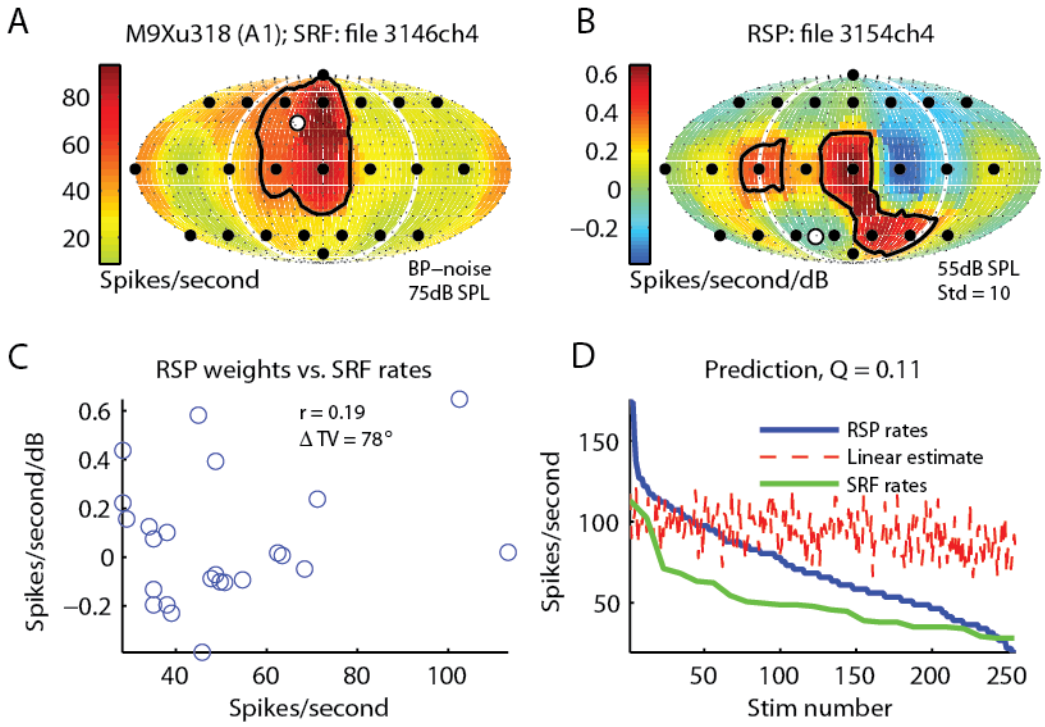


Figure 3.14.5. Comparison of 1- and 24-speaker receptive fields (5). An example neuron with poor agreement between the two conditions. (**a**) Spatial receptive field measured in the 1-speaker condition. (**b**) Spatial weighting function calculated using responses to random spatial profile (RSP) stimuli. (**c**) RSP weights vs. SRF weights. (**d**) Responses to RSP stimuli ordered from highest to lowest (blue), the linear prediction using the weighting function and the RSP level matrix (red), and the 1-speaker responses (green), ordered from highest to lowest.

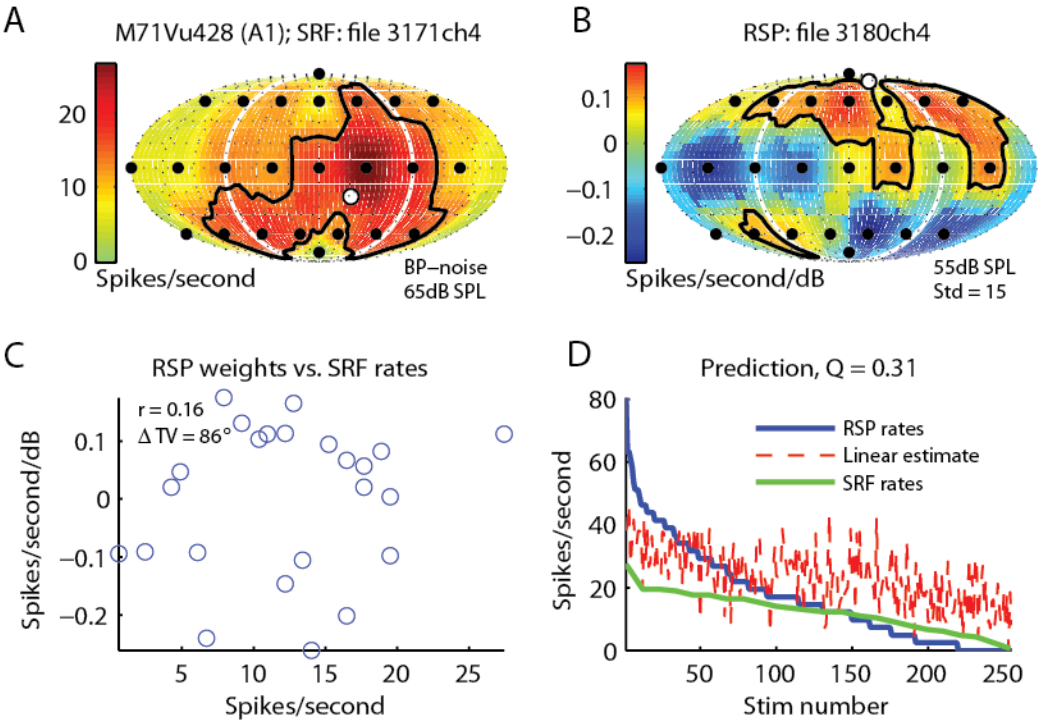


Figure 3.14.6. Comparison of 1- and 24-speaker receptive fields (6). An example neuron (2) with poor agreement between the two conditions. (**a**) Spatial receptive field measured in the 1-speaker condition. (**b**) Spatial weighting function calculated using responses to random spatial profile (RSP) stimuli. (**c**) RSP weights vs. SRF weights. (**d**) Responses to RSP stimuli ordered from highest to lowest (blue), the linear prediction using the weighting function and the RSP level matrix (red), and the 1-speaker responses (green), ordered from highest to lowest.

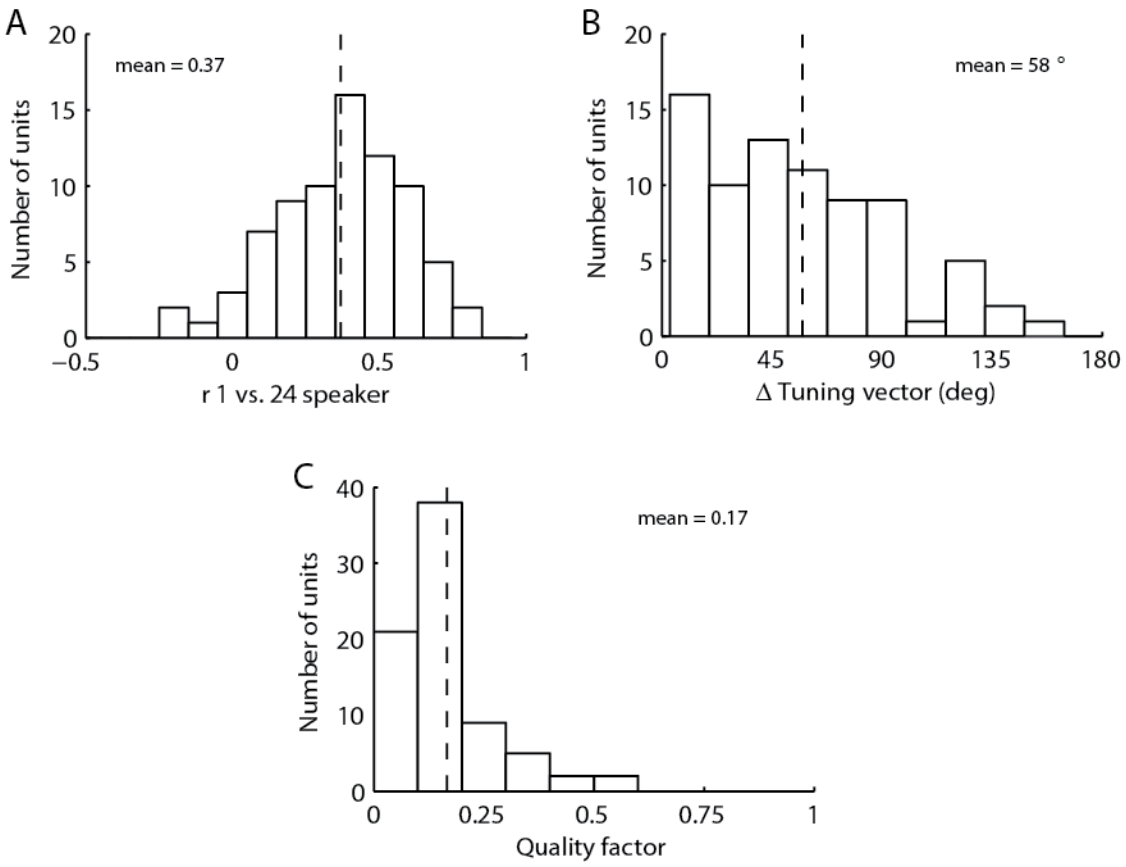


Figure 3.15. Distribution of 1- and 24-speaker agreement and prediction quality. (**a**) correlation coefficient and (**b**) tuning vector shift. (**c**) Prediction quality factor. Most units had very poor linear prediction performance. No effect of area on 1- and 24-speaker tuning agreement or linear prediction quality was observed.

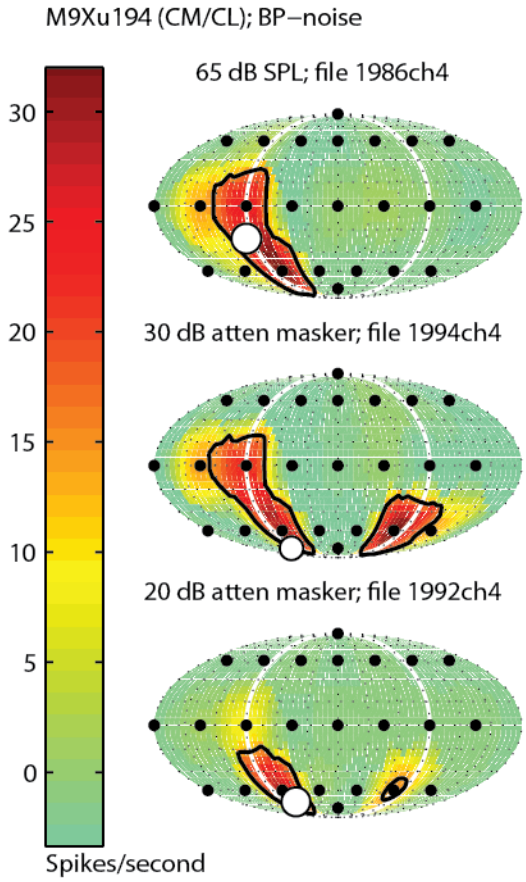


Figure 3.16.1. Spatial tuning in diffuse spatial nose (1). This unit shows relatively conserved SRF shape with a diffuse spatial noise. This unit's 1- and 24- receptive field agreement are shown in **Figure 3.14.4.**

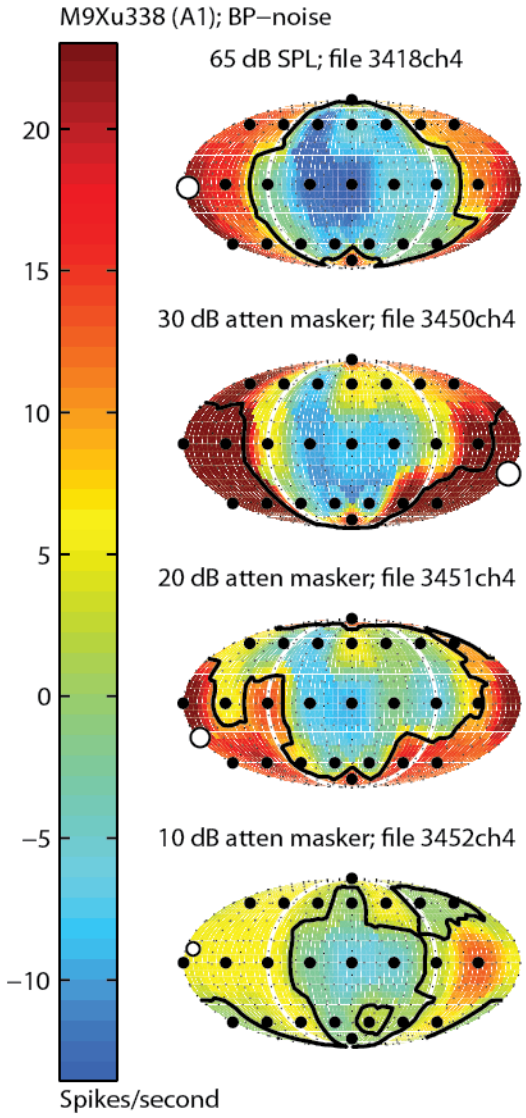


Figure 3.16.2. Spatial tuning in diffuse spatial nose (2). This unit shows highly conserved SRF shape with a diffuse spatial noise. 1- and 24- speaker correlation for this unit was 0.49.

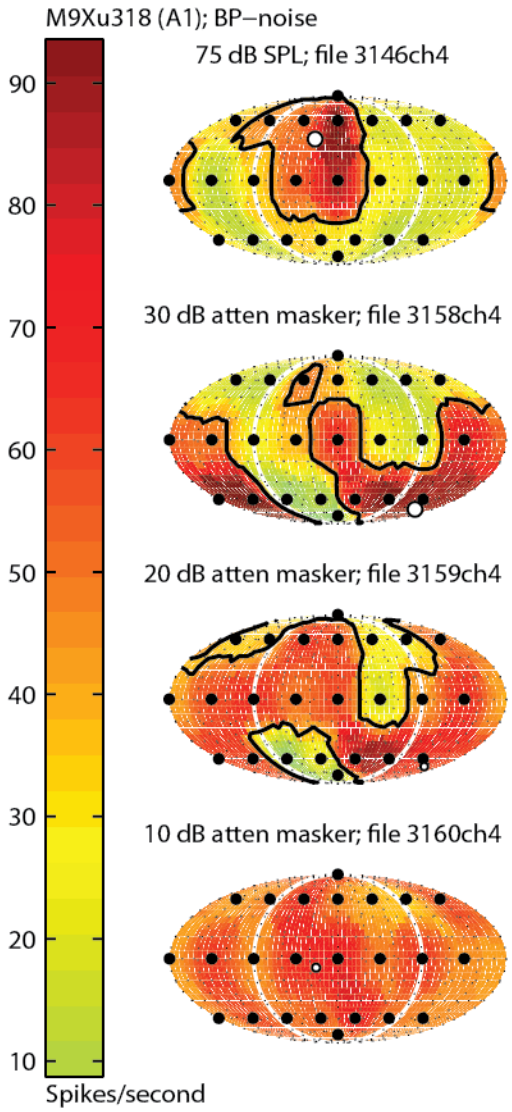


Figure 3.16.3. Spatial tuning in diffuse spatial nose (3). This unit shows poorly conserved SRF shape with a diffuse spatial noise. This units 1- and 24- speaker agreement are shown in **Figure 3.14.5.**

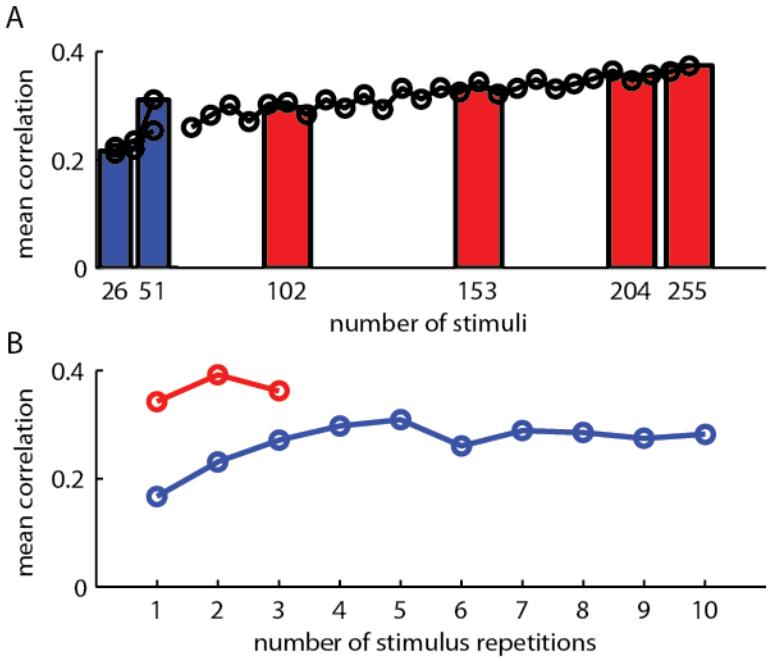


Figure 3.17. 1- and 24- speaker correlation vs. stimulus and repetition number. (**a**) 1- and 24- speaker correlation vs. the number of stimuli used to estimate the RSP weighting function. Blue bars represent neurons tested with a set of 51 stimuli and red bars represent neurons tested with a set of 255 stimuli (5 sets of 51). For neurons tested with 255 stimuli, the number of stimuli was increased 1 set of 51 at a time; the line represents permutations of 1-5 sets and the bars represent the average over all permutations of the same number of sets. 1- and 24- speaker correlation does not appear to saturate as stimulus number increases. (**b**). 1- and 24- speaker correlation appears to saturate near 3-4 repetitions for neurons tested with only 1 stimulus set.

Chapter 4: An auditory operant conditioning task for marmosets and measurement of minimum audible angle

4.1 Results

4.1.1 An auditory operant conditioning task for marmosets

4.1.1.1 Go/no-go task

We chose to implement a Go/No-Go type task suited for measuring detection and discrimination thresholds. The task is similar to those previously described for non-human primates (Brown et al. 1980; Sinnott et al. 1985). Figure 4.1 illustrates the behavior apparatus and paradigm. The objective in a Go/No-Go task is to respond to target sounds to receive reward while withholding responses when a target is not presented. Each behavior session is composed of a preset number of trials (typically 80-100), where each trial is composed of a variable duration 'inter-target interval' and a fixed duration 'response interval.' Inter-target interval duration is randomized between approximately 3 and 10 seconds but can be adjusted based on an animal's behavior (see Response Shaping section below); the response interval is dependent on the number and duration of targets but is typically 5 seconds in length. During an inter-target interval the subject hears either silence (in a detection task) or a series of background sounds (in a discrimination task). Behavioral responses during this time result in a mild punishment (see Response Shaping section below) and a restarting of the trial after the lick detector's infrared beam is clear for a preset duration. After the waiting period ends, target stimuli are alternated with background sounds during the response interval. The trial ends when the response interval has expired or a lick is detected during the response interval. Behavioral responses during this time are reinforced with approximately 0.1 – 0.2 ml food reward. During reward delivery, the program pauses to allow the

subject to consume the reward, beginning the next inter-target interval the after the lick detector's infrared beam is clear for a preset duration. If no response is detected, the next trial begins immediately.

A small percentage of trials are "catch trials," which are identical in length to target trials in their timing and structure but in which no targets are delivered (i.e., only silence or background sounds are heard during the response interval). Thus, during a catch trial the response interval is indistinguishable from the inter-target interval from the animal's perspective. A response during a catch trial response interval is referred to as a 'false positive' (or false alarm). The false positive rate gives a measure of response specificity from which an experimenter can create an adjusted hit rate or calculate (along with hit rate) a measure known as d' (Gescheider 1985) in order to determine an animal's perceptual sensitivity.

4.1.1.2 Response Shaping

The procedure to train subjects to perform behavior tasks is referred to as response shaping. This process is controlled by custom software in conjunction with monitoring by the researcher. After an animal has been adapted to sit in the restraining chair and accept food through the feeding tube, training proceeds through two phases. *Phase 1*: food rewards are delivered following an auditory stimulus such as a white noise or pure tone while the animal's behavior is monitored via closed circuit television and software. In this phase reward is not contingent on the subject's behavior response. Animals soon start to associate the sound with food reward and begin showing anticipatory licking responses. *Phase 2*: reward delivery is made to be contingent on licking to the conditioning sound. The animal stays in phase two until the hit rate is consistently above 80% and the false positive rate is consistently lower than 25%. The animal is then considered trained, and testing on a detection task begins (for example, to determine hearing thresholds). Alternatively, animals can then be moved to a more complex discrimination task in which silent periods in the

inter-target interval are replaced with audible background sounds. Where detection tasks are typically used to probe an animal's hearing sensitivity, discrimination tasks are more generally used to test an animal's ability to perceptually separate two sounds along some dimension.

Because the animal has been trained to lick in response to sounds, the presentation of audible background sounds during a discrimination task will usually bring a strong response from the subject at first. For this reason, it can be helpful to continue presentation of background sounds without pausing in response to licks until no licks have been recorded for several seconds, and then present the first target. Often, a monkey responds to this first target (provided it is easy to distinguish from the background) and continues to respond to further discrimination targets. Then, the process of false positive reduction repeats again until below the nominal level of 25%.

In some animals extra care is taken to reduce false positives. Any observer with some amount of internal noise will produce false positives, the probability of which is controlled by the response bias. In order to shift response bias and reduce false positives, several methods are employed, depending on each animal's propensity to lick in error. One way to reduce false positives is to reduce the target probability (Gescheider 1985), which can be achieved by increasing the inter-target interval length or the frequency of catch trials. Additionally, the number of targets below the perceptual threshold of the animal can be decreased. This doesn't reduce stimulus probability *per se*, but rather reduces the number of targets for which a guess will result in a reward. For most animals, a ratio of response window length to inter-target interval length of about 0.5, less than 25% undetectable targets, and 20-30% sham trials is sufficient to keep false positives to an acceptable level.

For some animals, introducing an additional mild punishment for errors is helpful in reduction of guessing behavior, particularly early in training. We have used the following: (1) the inter-target interval is re-started and lengthened, (2) a "timeout" period (as described previously)

is introduced , (3) the timeout is accompanied by a temporary shutting off of the chamber house light (blackout) and (4) the timeout period is accompanied by a puff of air delivered to the animal's back or tail. For most animals, a timeout is sufficient to reduce false positives to acceptable levels.

4.1.1.3 Performance in a detection task

To quantify task learning and performance, we trained five common marmosets (two male, three female) between two and five years of age on a Go/No-Go detection task. After marmosets became adapted to the restraint chair and first displayed anticipatory licking to sounds (*Phase 1,* Response Shaping), we quantified learning behavior through *Phase 2* of training. Hit rates increased and false positives decreased as the animal learned to associate sound with food reward, and training was considered complete when 4 of 5 consecutive sessions had been completed with at least 80% hit rate and less than 25% false positives. The "time to train" for a particular animal was the first session of *Phase 2* in which the subject reached this criterion of the 4 required. For 2 marmosets the detection sound was a 6 kHz pure tone, and for the other 3 the sound was a broad band noise token band-pass filtered between 2 and 32 kHz. These stimuli were chosen for the purposes of future psychophysical testing: the first group was later tested for pure tone detection thresholds (Osmanski & Wang 2011), and the second group for spatial hearing acuity. Average time to train across all animals was 12 sessions with a standard deviation of 6 sessions. Figure 4.2 shows *Phase 2* learning curves for 5 animals trained over 2 to 3 weeks.

Figure 4.3 illustrates the time course of licking behavior and shows response latency and licking duration distributions for a representative behavior session. Response latency was measured as the elapsed time from the onset of the first target stimulus to the first lick. Licking duration was measured as the time from the first lick to the offset of the last lick. Sessions lasted 80 to 100 trials (30% of which were sham trials), after which there was a tendency for a reduction

in motivation, likely due to animals becoming sated. The average session duration across all subjects at the end of the training period (last 5 sessions) was 32 minutes with a standard deviation of 6 minutes.

4.1.1.4 Application to electrophysiology

A crucial goal of the behavior design was to allow the pairing of auditory perceptual tasks with single unit neurophysiology; the behavior setup was therefore designed specifically to be compatible with our neural recording methods. The setup (Figure 4.1) utilizes a modified version of the restraining chair used in our previous studies, allowing electrophysiology recordings to be performed as normal. However, it is important that licking, which results in jaw and tongue movement, does not adversely affect electrode stability or electrical signal strength. To show that single-unit recordings are possible during licking, we trained an implanted, head-fixed marmoset to discriminate sound source locations while recording single-unit responses during task performance. Although this task results in muscle movement of the jaw and tongue, as well as the presence of an electronic device in the vicinity of the recording equipment, there was no obvious reduction in recording stability or electrical signal quality. Figure 4.4 shows a filtered voltage signal from an electrode recording single unit activity in marmoset auditory cortex during task performance. There is no appreciable movement or electrical artifact before or after lick detection, even though the animal's jaw and tongue are active during these times.

4.1.2 Perceptual acuity for sounds varying in azimuthal and elevation

We used this behavioral task to measure minimum audible angle in marmosets along the horizontal and vertical dimensions. The 5 subjects used in this study quickly generalized to the location discrimination task from the detection task, after which they were tested daily until criterion was met (see Chapter 2, section 10). Not all animals were tested in each condition. Sound level was roved ±10dB to avoid the use of absolute level as a cue. The chair used in this

portion of this experiment was a wire mesh chair (the same chair as chapter 3 without the electrophysiology setup) designed to reduce acoustic occlusion and reflection, and is shown along with the speaker arrangement in figure 4.5.

4.1.2.1 Horizontal acuity: Gaussian noise

For all noise stimuli, noise tokens were generated uniquely for every trial. The ability of marmosets to discriminate sound source azimuth is illustrated in figure 4.6. Typical performance variability is shown using psychometric curves in figure 4.6a. All following psychometric functions are shown as averages with standard error of mean over several sessions. The 5 animals tested showed generally good agreement in thresholds, with a mean threshold of 15° and standard deviation of 4°. Most of the variation among the subjects occurred at 15° separation; most could reliably discriminate 22.5°, yet none could do so for 7.5° separation.

4.1.2.2 Horizontal acuity: random spectral shape stimuli

Although sound localization in the horizontal plane relies primarily on binaural cues, HRTF shape has also been observed to change with azimuth in several species (Wightman & Kistler 1989; Rice et al. 1992), although these changes are mostly confined to higher frequency cues in marmosets (Slee & Young 2010), and cats have been shown to not depend heavily on specific frequency cues in a horizontal discrimination task (Huang & May 1996b). Still, there is a possibility that in addition to binaural cues, monaural spectral cues could be used to perform a horizontal discrimination task (Butler 1986; Van Wanrooij & Van Opstal 2004). To reduce the possibility that marmosets could be using spectral cues for azimuth discrimination, we tested three marmosets' ability to discriminate the location of a sound which varied spectrally on each stimulus presentation. This would make it difficult to compare successive stimuli on the basis of frequency spectrum. We used a stimulus created for the measurement of spectral weighting functions, called random spectral shape, or RSS stimuli (Yu & Young 2000). We hypothesized

that if subjects were relying on spectral cues to perform azimuth discrimination, thresholds should be higher when discriminating RSS stimuli. Psychometric functions for RSS stimuli are shown in figure 4.7. Thresholds were as low, if not lower, than in the Gaussian noise condition (mean 13°), suggesting that the subjects used binaural cues to discriminate horizontal locations.

4.1.2.3 Horizontal acuity: rear locations

Binaural cues are often weaker at rear locations (Harrison & Downey 1970), and data suggest that localization is less accurate at rear locations compared with frontal locations (Oldfield & Parker 1984). However, as marmosets are a tropical arboreal species living in a highly visually occluded environment, we thought it would be interesting to test whether they possessed relatively heightened localization abilities outside the visual field. We tested several marmosets' ability to discriminate azimuth at rear locations using both Gaussian noise and RSS stimuli; these results are summarized in figure 4.8. Animals were not tested as extensively in this condition as the front location; however thresholds appeared to be generally elevated when compared with the front locations, indicating that, at least without additional training, rear acuity in marmosets is worse than frontal acuity.

4.1.2.4 Vertical acuity

While horizontal spatial information is contained in the differences in the signals between the two ears, additional information is required to compute sound source location in 3-dimensional space. To do this, the auditory system makes use of spectral cues generated primarily by the pinna. These cues are typically located at higher frequencies. We measured vertical location acuity first using the same Gaussian noise stimuli used to test horizontal acuity. At first, discrimination of vertical locations was difficult for the marmosets. Two marmosets which initially reached the criterion set for discrimination thresholds (and had quite stable performance) had apparently extremely large vertical minimum audible angles (59° and 53°). We thought it unlikely that

spectral cues provided by the ears in marmosets could afford only coarse discrimination in the vertical axes, and continued to test subjects on the task. Gradually, performance increased, and after criterion was again met, thresholds were again calculated. Psychometric functions for three animals after this extended training are shown in figure 4.9; average threshold was 17°.

The HRTF of mammals contains several components which could be useful for directional hearing. These include a "mid-frequency" component, in which there exists a prominent spectral notch (sometimes called the "first notch") which varies in frequency in a somewhat predictable manner with elevation, and a "high frequency" component, which inhabits the upper range of audibility and contains cues which can vary in a less predictable manner (Rice et al. 1992; Slee & Young 2010). It has been shown previously in cats that accurate *absolute* localization (that is localization identification by head orientation) is dependent on mid-frequency cues, degrading when only high frequency cues are available (Huang & May 1996a), while the lack of mid-frequency cues did not substantially increase thresholds for vertical discrimination (Huang & May 1996b). In marmosets, this first notch has been shown to vary between 12 and 24 kHz (Slee & Young 2010). To reduce the possibility that marmosets were using high frequency cues, vertical discrimination acuity was tested in several animals using Gaussian noise filtered between 4 and 26 kHz (figure 4.10, adapted from Slee and Young, 2010). Psychometric functions for five marmosets discriminating elevation using the mid-frequency Gaussian noise are shown in figure 4.11. The average threshold was 22°, slightly higher than when the stimulus included higher frequency information up to 32 kHz, but the difference was not significant ($p = 0.29$). Three animals were tested using both stimuli. Two of the animals displayed higher thresholds in the mid-frequency case (M3T: 19° vs. 26°; M71V: 12° vs. 18°), whereas one had slightly lower thresholds (M94W: 20° vs. 19°). We additionally tested two animals using sounds which were filtered to only include stimulus information below the location of the first spectral notch. Animals tested with this stimulus (Gaussian noise, 4 – 12 kHz) exhibited extreme difficulty in

discriminating vertical location; one animal had a threshold of 72°, while another could not discriminate any of the locations in the frontal hemifield. Both animals, however, could reliably discriminate front and rear locations (figure 4.12).

4.2 Discussion

4.2.1 Go/no-go licking task: comparison with other behavior methods

Limited marmoset psychoacoustic data from an early study was collected using negative reinforcement (i.e. shock avoidance; Seiden 1957). Assuming that positive reinforcement would be more advantageous for both behavioral and physiological studies, our lab tested several food-reward protocols over the past decade. In addition to the licking strategy described here, our lab has investigated both lever manipulation (the author and previously Dennis Barbour) and eye position tracking (Simil Roupe, Elias Issa, David Kim, and Poppy Crum). Behavioral reporting via lever movement seems a logical choice, as it allows the reporting apparatus to be located far from the head and ears while potentially allowing for multiple response types (e.g. a left vs. right lever movement). Eye tracking has similar advantages: several saccade targets can be used, and equipment is out of the way, provided the high-speed camera can be positioned such that the acoustic field is not disturbed. Others in the lab had some early success with eye tracking, but there was very little success with the lever. Lack of success with the lever task may have been due to the physically constraining marmoset chair. While experimenting with lever training, however, marmosets were apt to lick at the feeding tube after a conditioning stimulus. In one telling case, a marmoset which was being trained to pull on a manipulandum to obtain juice reward (not contingent on any target sound) never pulled on its own but very quickly began licking as soon as the manipulandum was moved by some external means. The tendency to lick to acquire food may be related to feeding patterns of marmosets in the wild, which include chewing holes in tree bark to feed on exudate (Bouchardet da Fonesca & Lacher 1984). Alternately, it could simply be that it is easier to train an action which is already necessary for food intake (marmosets must lick to ingest the reward regardless of whether reward delivery is contingent upon licking).

There are two potential disadvantages of lick reporting: first, the lick detector as described here has only one reporting option, ruling out a multiple forced choice task. The setup could be amended by adding a second feeding tube and lick detector, but this would be more difficult in a head-fixed neural recording setup. Second, as the behavior apparatus is near the head and ears, possible acoustic field distortions should be considered. This issue would need to be addressed when conducting studies of spatial hearing, but it is possible to drastically reduce the amount of material holding the LED and phototransistor in place (for example by utilizing coiled wire, as was done here). We believe that these drawbacks are far outweighed by the relative simplicity of training marmosets in the licking task.

4.2.2 Measurement of sound location discrimination in the common marmoset

As a tropical arboreal species, marmosets need to navigate their environment using acoustic spatial cues. Spatial processing is therefore an important function performed by the marmoset's auditory system. Not surprisingly, acoustic measurements indicate that marmosets possess cues to localize sound sources in azimuth and elevation (Slee & Young 2010). We therefore applied the task to investigate the ability of marmosets to discriminate sound location.

The average measured threshold for horizontal localization of 15° (13° for RSS stimuli) is higher than many other mammals tested. Several species have been shown to discriminate sound locations separated by less than 10° in azimuth, including cats (Heffner & Heffner 1988a), macaques (Brown et al. 1980), and opossums (Ravizza et al. 1972). Despite significant training, 10° seemed to be unachievable for azimuth discrimination by marmosets in the current study. It is unlikely that this limit was due to a lack of motivation, as false positive rates were not particularly low; all animals had false positive rates higher than threshold on several occasions. Another possibility is the use of sounds filtered above 2 kHz, which is above the frequency region where ITD cues are most useful (Stevens & Newman 1936). In humans, the upper limit for ITD cues has

been measured at 1.3 kHz (Klumpp & Eady 1956). We believe this is also an unlikely limiting factor, as cats and macaques have both been shown to exhibit high horizontal acuity using test stimuli of comparable spectra (Brown et al. 1980; Huang & May 1996b). Also, there is some evidence to suggest that several small mammals can use interaural timing cues at frequencies significantly higher than 2 kHz (Masterton et al. 1975; Heffner & Masterton 1980; Heffner & Heffner 1985). This is possibly to compensate for the lack of strong ILD cues in the middle frequency region due to small head size and also because ITD cues in these species become ambiguous at higher frequencies than animals with larger heads. ITD ambiguities are the result of the maximum producible ITD at a given frequency being larger than the period of the stimulus, leading to a non-affine mapping between ITD and lateral position for those frequencies. In fact, fine structure phase locking in the auditory nerve has been observed at frequencies up to 4 kHz in both cats and squirrel monkeys (Johnson 1980; Rose & Brugge 1967). In marmosets, ILD cues are relatively small below 5 kHz (Slee & Young 2010). We believe a better explanation for the high thresholds observed here in marmosets compared to those in cats and macaques therefore is the relatively small head size of marmosets. As binaural cues are known to be dependent on the physical distance between the two ears (for ITD) and the size of the head (for ILD), it is perhaps not surprising that marmosets do not appear to be expert localizers. In a meta-analysis, there was shown to be a relatively good correlation between head size (as defined by maximum interaural time difference) and horizontal discrimination thresholds (Brown & May 2005). Placing marmosets into this dataset suggests that the performance measured is almost exactly what would be predicted based on head size (figure 4.12).

Although sound localization in the horizontal plane relies primarily on binaural cues, HRTF shape has also been observed to change with azimuth in several species (Wightman & Kistler 1989; Rice et al. 1992), and monaural spectral cues can be used to perform a horizontal discrimination task (Butler 1986; Van Wanrooij & Van Opstal 2004). As this task measured

discrimination ability rather than absolute localization accuracy, the usefulness of these spectral cues could be increased. RSS thresholds which were equally low or lower than in the Gaussian noise condition suggest that the marmosets used primarily or exclusively binaural cues to discriminate horizontal locations.

Although binaural cues are often weaker at rear locations (Harrison & Downey 1970), and data suggest that localization is less accurate at rear locations compared with frontal locations (Oldfield & Parker 1984), an animal suited to an arboreal environment might possess heightened sensitivity to space outside the visual field. The observations that rear acuity was lower when compared with front locations indicated that this may not be the case for marmosets.

Subjects' thresholds for vertical localization were higher than those for horizontal localization. This finding is consistent with at several species previously tested, such as the chinchilla (Heffner et al. 1995) and the opossum (Ravizza et al. 1972), but vertical and horizontal acuity have been shown to be roughly equal in several species, such as cats (Martin & Webster 1987) and macaques (Brown et al. 1982). The exceptional vertical discrimination in these species, however, is degraded by the removal of high frequency energy (Brown et al. 1982; Huang & May 1996b). The HRTF of mammals includes a "mid-frequency" component, in which there exists a prominent spectral notch (sometimes called the "first notch") which varies in frequency in a somewhat predictable manner with elevation, and a "high frequency" component, which inhabits the upper range of audibility and contains cues which can vary in a less predictable manner (Rice et al. 1992; Slee & Young 2010). In cats, accurate *absolute* localization (that is localization identification by head orientation) is dependent on mid-frequency cues, degrading when only high frequency cues are available (Huang & May 1996a), thus vertical discrimination thresholds using stimuli with very high stimulus energy may overestimate the actual localization ability of listeners. In cats discriminating "mid-frequency" sounds (5-18 kHz), vertical localization thresholds were shown to be higher than azimuth thresholds (Huang & May 1996b).

In marmosets, this first notch has been measured and varies between 12 and 24 kHz (Slee & Young 2010). The initial test stimulus, with a frequency range of 2-32 kHz, did include energy in marmosets' "high frequency" region. To test acuity without access to information at high frequency cues outside the range of the first notch, vertical discrimination acuity was tested in several animals using Gaussian noise filtered between 4 and 26 kHz. The average threshold of 22° was slightly higher than when the stimulus included higher frequency information, consistent with previous findings in cats (Huang & May 1996b). Finally, to test whether marmosets were in fact using energy in the notch region and not using low frequency cues (lower than 12 kHz), we tested acuity in 2 marmosets using a 4-12 kHz Gaussian noise. Performance was very poor using this stimulus, indicating that subjects did not use low frequency cues to perform vertical location discrimination.

A final consideration is the amount of training marmosets required to reach criterion for vertical discrimination thresholds. Although long training periods have been required to reach asymptotic performance in tasks involving spatial cues in lab animals in previous studies (e.g. Scott et al. 2007), such poor thresholds were not observed initially in azimuth discrimination. One possible explanation is that marmosets are not particularly well suited to performing location discrimination in the median plane, as spectral cues have been shown to be less regular there than at lateral positions (Slee & Young 2010). However, it is also possible that the laboratory housing environment is not optimal for the development of such abilities. At least for azimuth, it has been shown that environmental acoustic factors can play both negative (Efrati & Gutfreund 2011; Pan et al. 2011) and positive roles (Cai et al. 2009) on the development of perceptual abilities. One study showed that in animals with environmentally degraded processing, perceptual abilities could be normalized by behavioral training (Pan et al. 2011). It may be interesting to compare coding for elevation between trained and untrained marmosets.

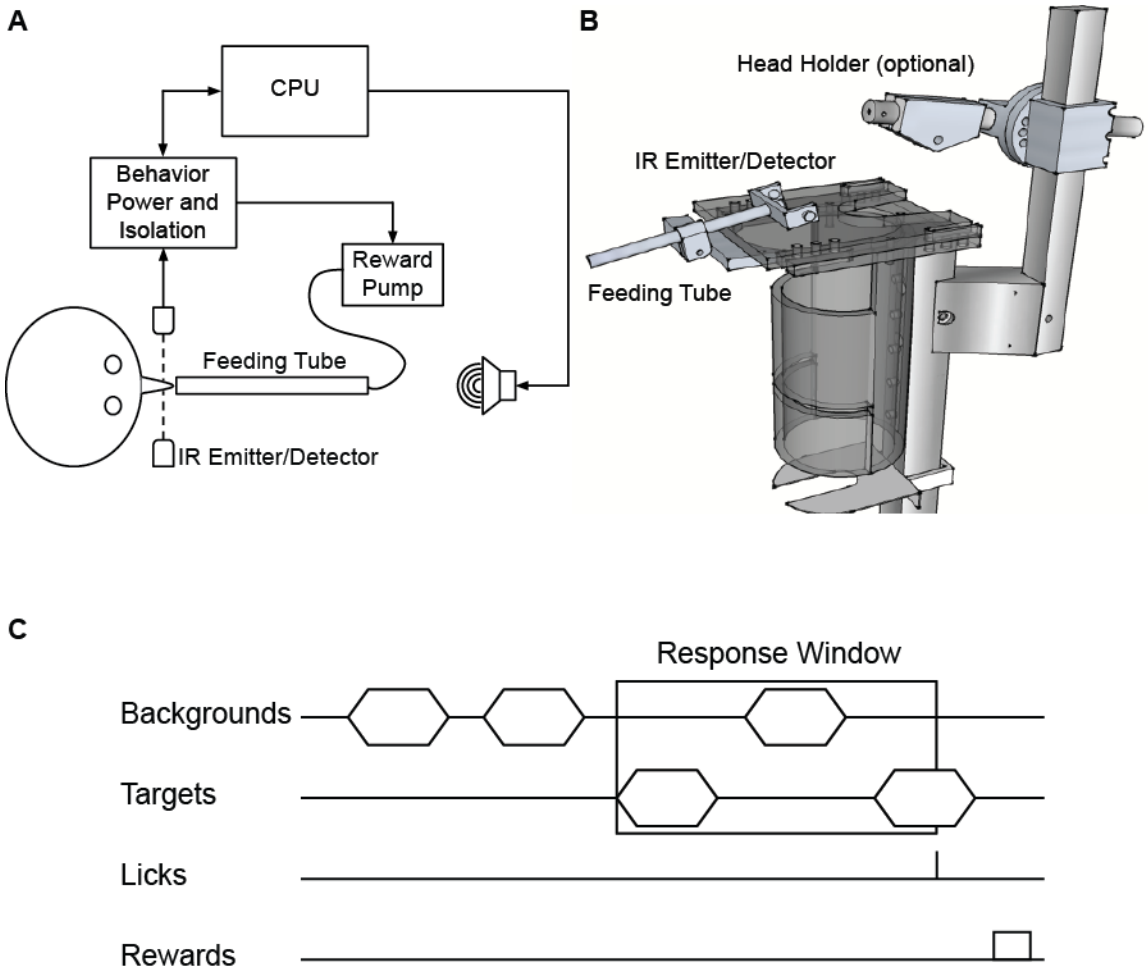


Figure 4.1. Behavior apparatus and paradigm. (**a**) Marmoset chair with feeding tube, infrared lick detector, and optional head restraint mechanism for single-unit recording. The neck plate slides out to allow a marmoset to enter the chair from below. After securing neck plate, the feeding tube can be adjusted to create a comfortable reach for each monkey. (**b**) Schematic of task setup. Sounds are played from free field speakers while marmosets lick to target sounds for a reward which is delivered by a syringe pump via a feeding tube. Lick responses are recorded when the infrared beam is broken by the animal's face or tongue. Behavior apparatus are controlled by a personal computer and powered by a custom built power supply and electrical isolation module. (**c**) After a variable number of background stimuli (or silent periods, for the detection task), targets will begin alternating with the background stimuli/silent periods. If a lick is registered within the preset number of alternations, a food reward is given. After the animal has finished consuming the reward (as measured via the lick detector), the next inter-target interval begins with background stimuli or silent intervals. A lick outside of a target interval results in a timeout.

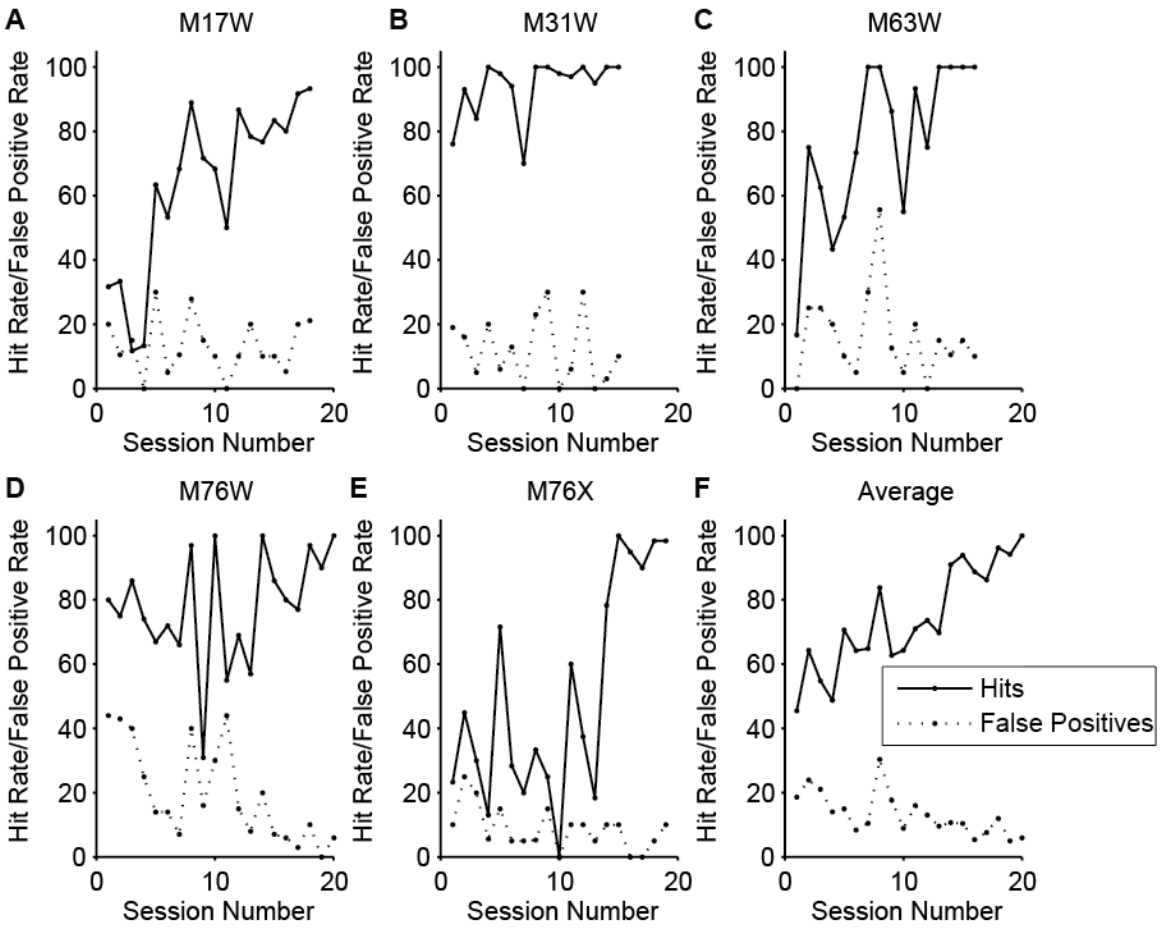


Figure 4.2 Detection task learning curves. (**a-b**) Learning curves for 5 naive marmosets performing an auditory detection task with broad band noise or pure tone stimuli. Data represent training *Phase 2* (see Response Shaping). Training is considered complete when 4 of 5 consecutive sessions have been completed with at least 80% hit rate and less than 25% false positives. Average time to train across all animals was 12 with a standard deviation of 6 sessions. (**f**) Average hit and false positive rates over all training sessions. Later sessions had fewer data points averaged due to some animals completing training more quickly than others.

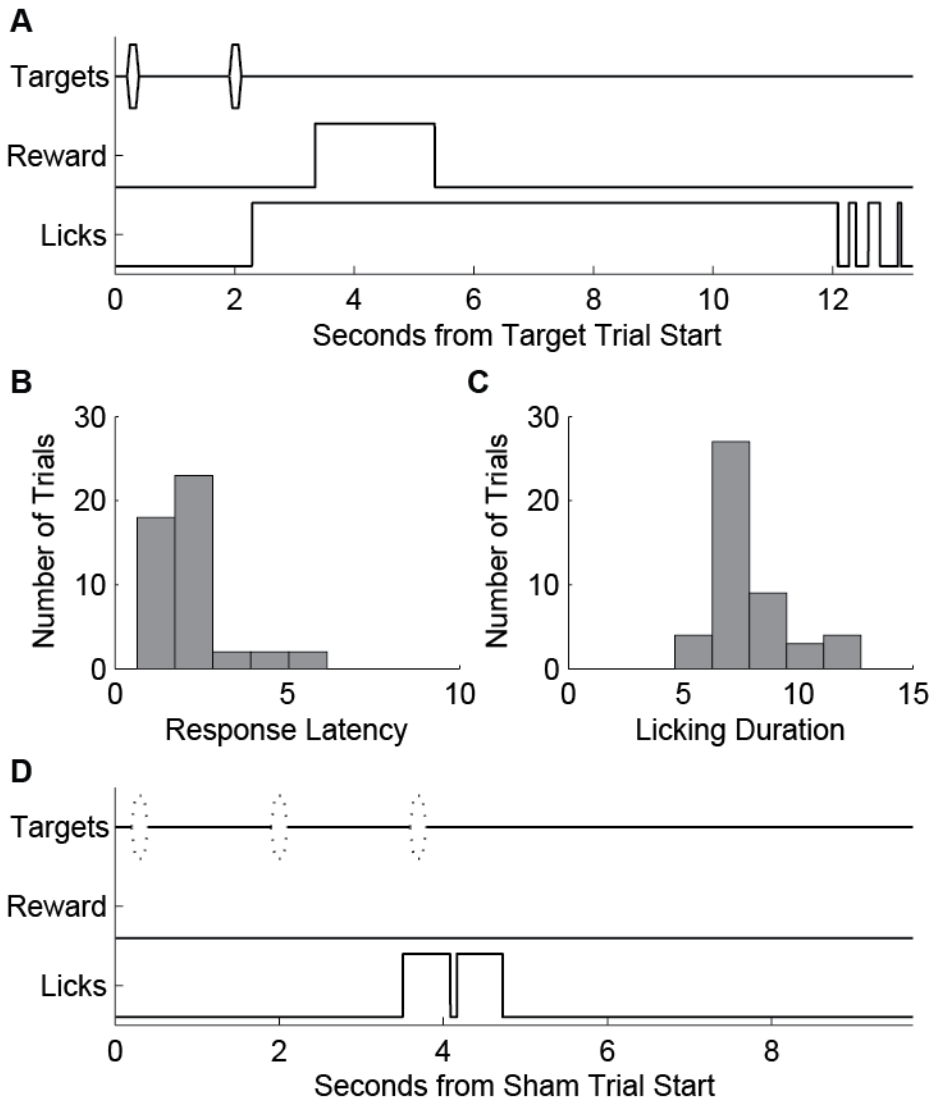


Figure 4.3. Licking behavior. (**a**) Example of a licking response to a target trial along with reward and feeding behavior to target trials for a representative behavior session. (**b,c**) Distribution of response latencies within the same session (**b**), measured as the elapsed time from the onset of the first target stimulus to the first lick, and lick durations (**c**), measured as the time from the first lick to the offset of the last lick. (**d**) Example sham trial with an error response. Sessions lasted 80 to 100 trials of which 30% were sham trials.

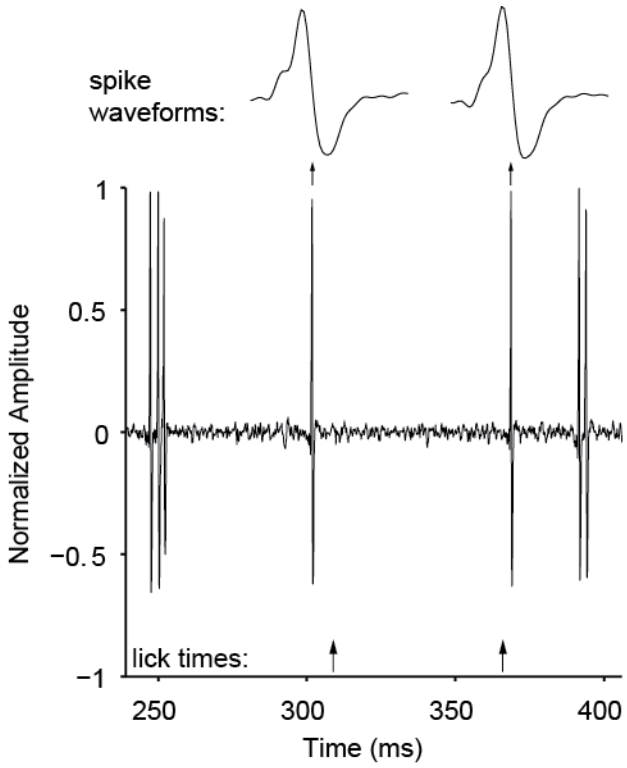


Figure 4.4. Single unit recording during behavior. Example of voltage signal, high pass filtered for spike sorting, from a high impedance microelectrode recording single unit activity in marmoset auditory cortex during task performance. Time is referenced to pre-stimulus delivery interval. The licking behavior can be performed without compromising recording stability (meaning that units can be held reliably) or signal quality. Note that spikes can be easily discerned both before and after a lick is detected.

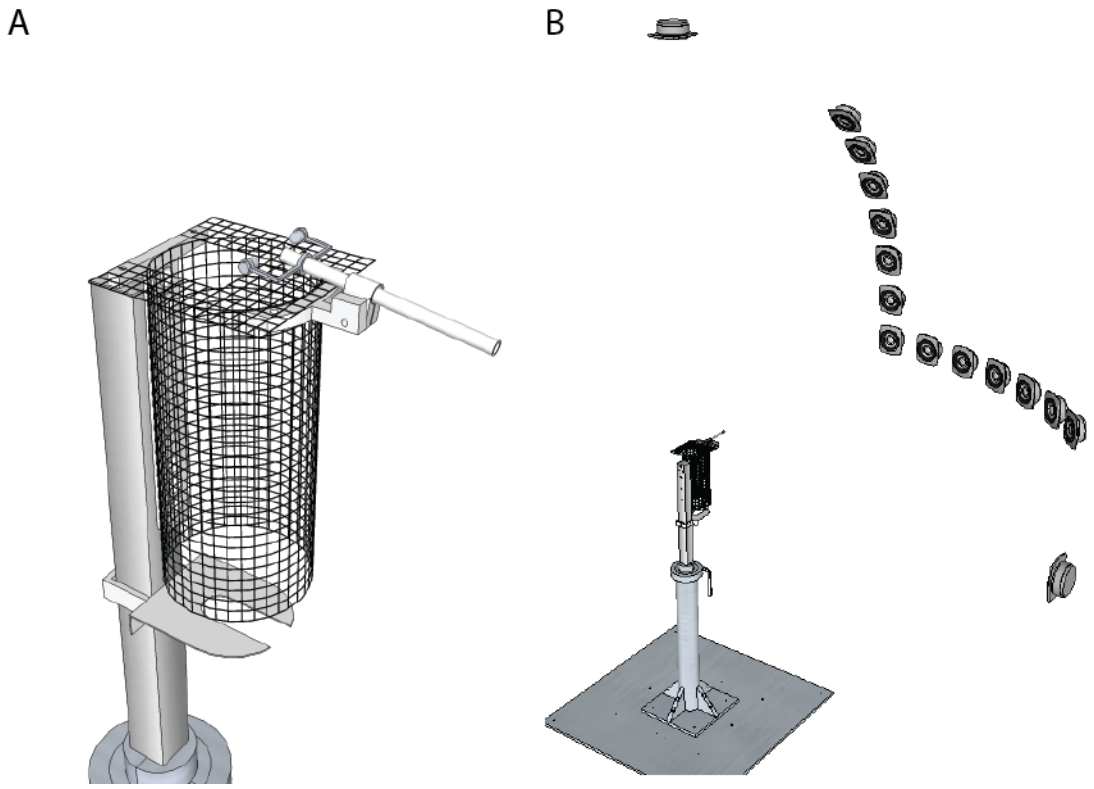


Figure 4.5. Wire chair and speaker arrangement for MAA. Behavior sessions for minimum audible angle measurements were conducted while marmosets were seated in a steel wire chair (**a**) built to reduce acoustic reflections. (**b**) Localization discrimination thresholds were measured in three conditions: frontal azimuth, rear azimuth, and vertically along the median plane. For rear azimuth discrimination, the chair was rotated 180°. Head position was monitored with a closed circuit camera system and custom image processing software (Matlab). Marmosets were required to have heads facing forward in order for behavioral trials to proceed. Rear location is not pictured.

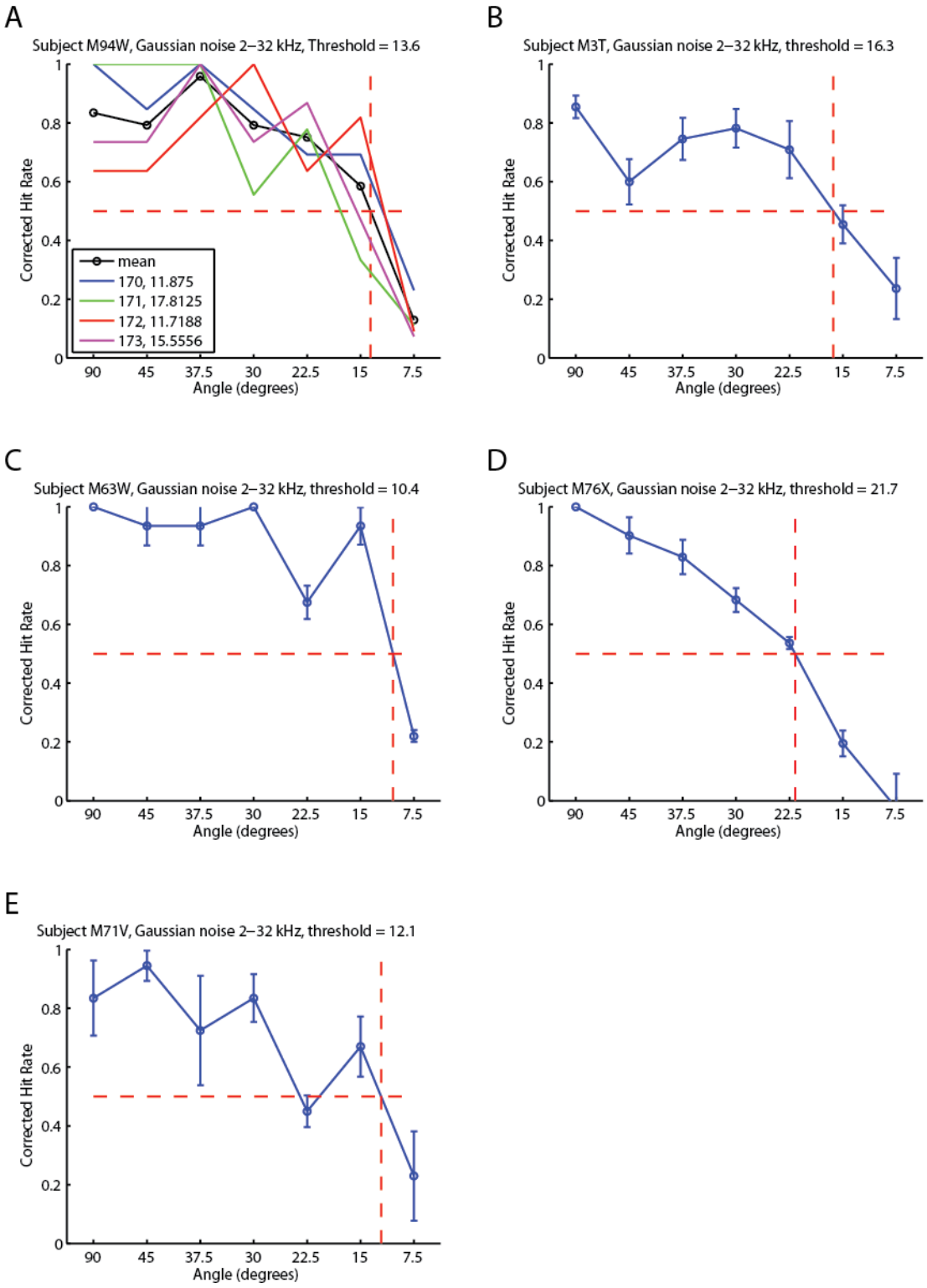


Figure 4.6. Frontal azimuth discrimination: Gaussian noise. (**a-e**) Psychometric functions for marmosets detecting changes in frontal azimuth location. Individual sessions are shown for one animal in (**a**).

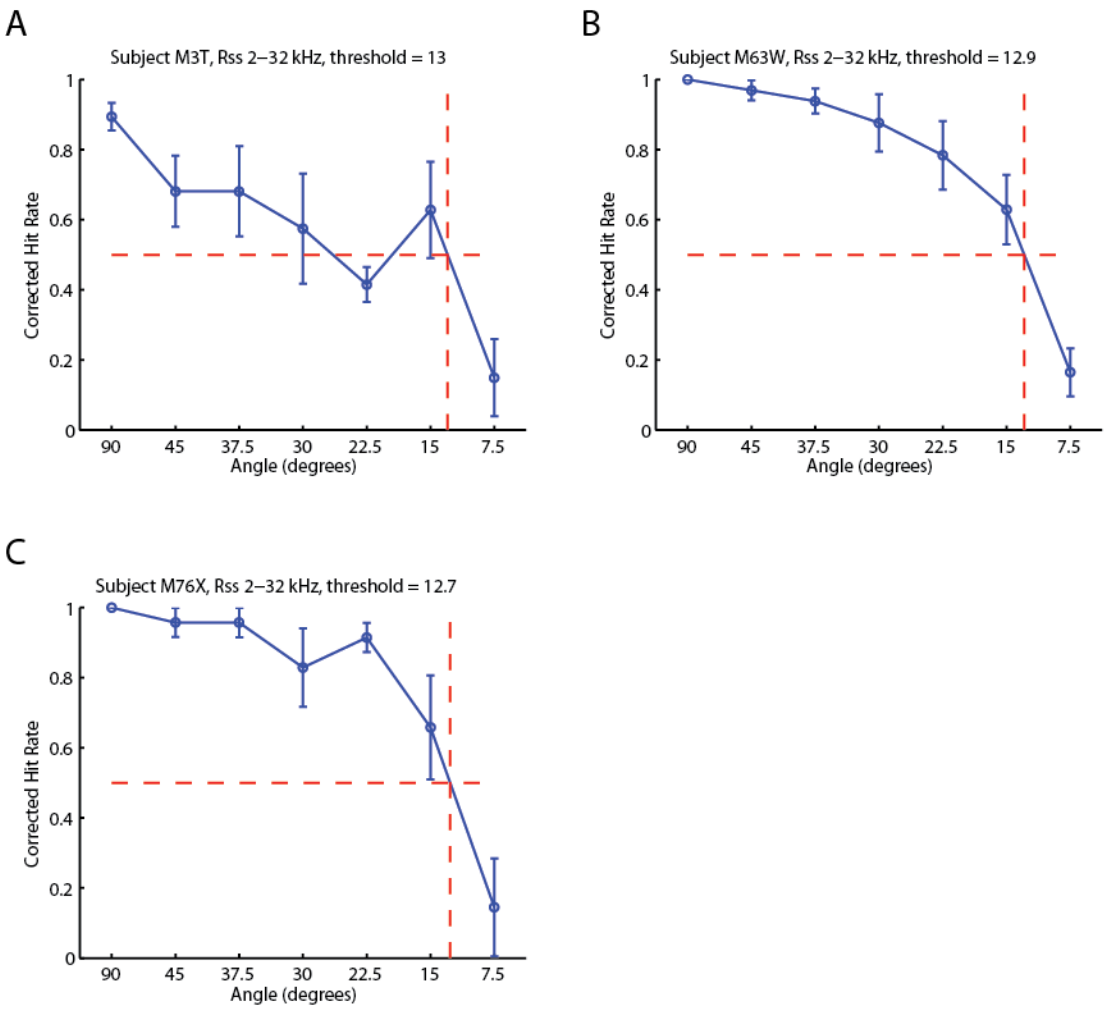


Figure 4.7. Frontal azimuth discrimination: RSS. (**a-c**) Psychometric functions for marmosets detecting changes in frontal azimuth location where the test stimulus varied in spectral shape from stimulus to stimulus. Thresholds for 2 of 3 animals are lower than those in the Gaussian noise condition.

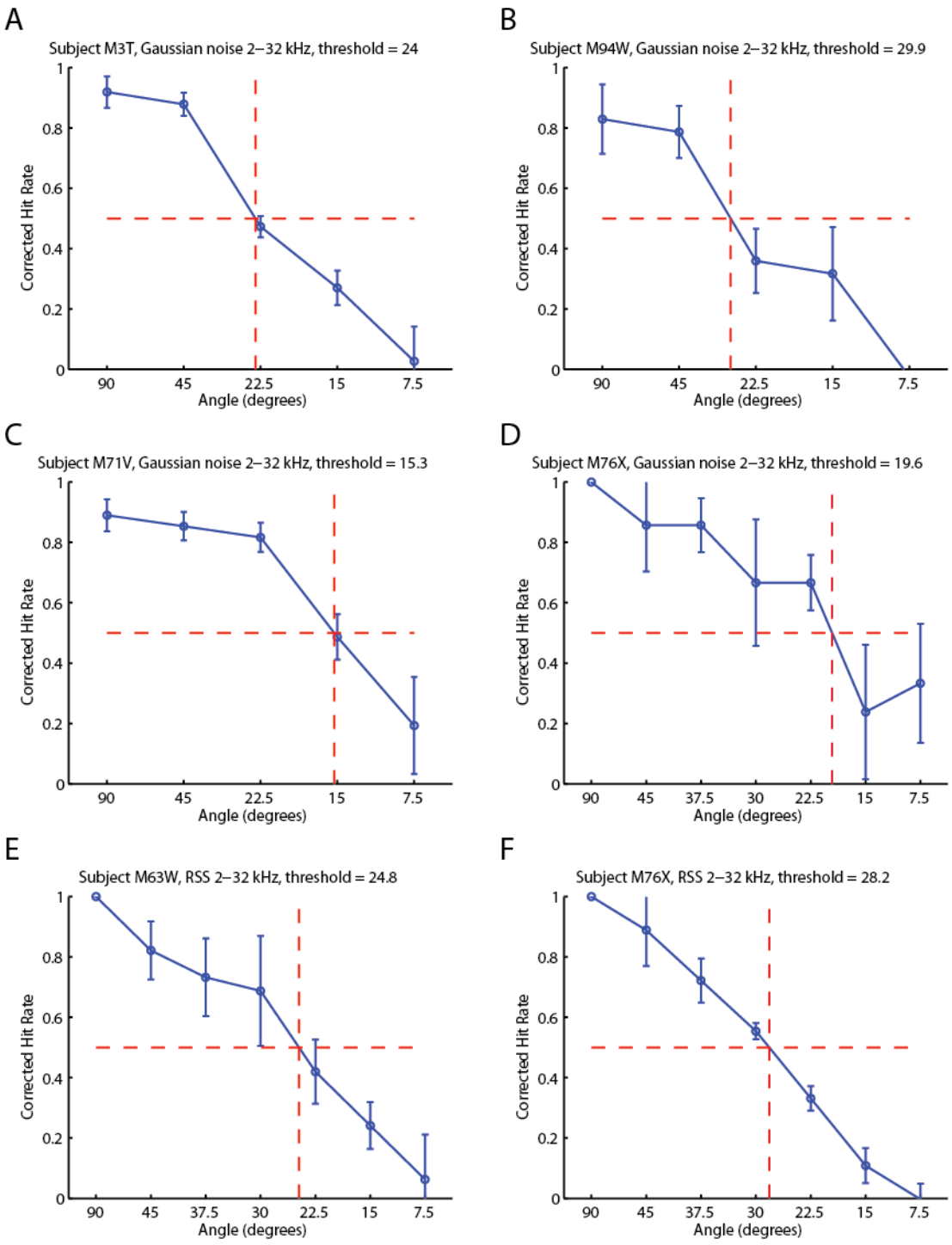


Figure 4.8. Rear azimuth discrimination: Gaussian noise and RSS. (**a-d**) Psychometric functions for marmosets detecting changes in frontal azimuth location where the test stimulus was a Gaussian noise. (**d,f**) Psychometric functions for rear discrimination using RSS tokens.

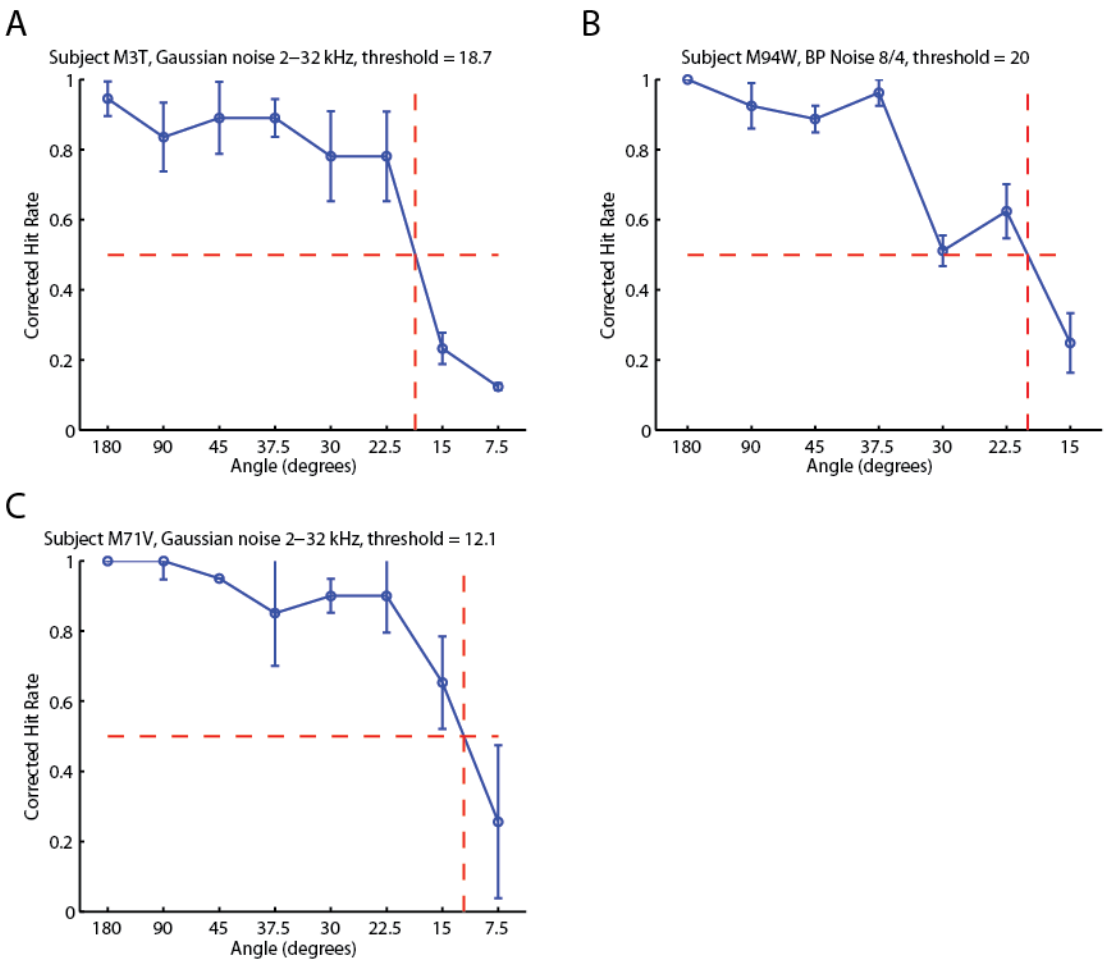


Figure 4.9. Frontal elevation discrimination: 2-32 kHz Gaussian noise. (**a-c**) Psychometric functions for marmosets detecting changes in frontal elevation location where the test stimulus was a Gaussian noise with energy above the first spectral notch region.

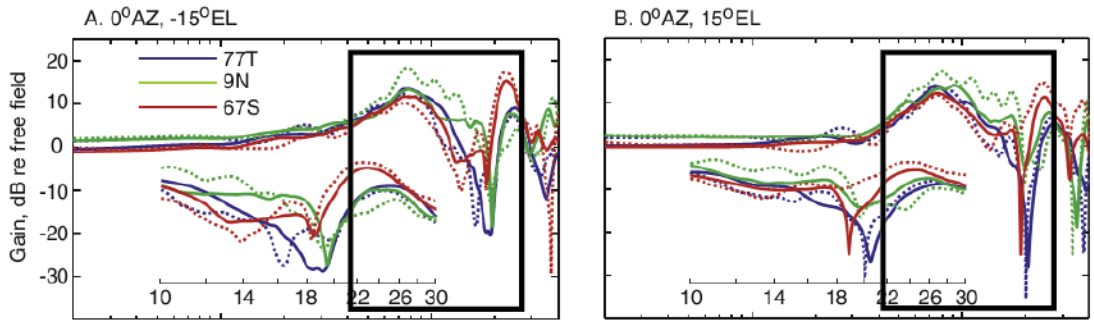


Figure 4.10 Marmoset HRTF. The spectral notches in the marmoset have been previously shown to exist in the 12-26 kHz frequency range. To reduce the possibility that marmosets were using high frequency cues outside this range, vertical discrimination acuity was tested in several animals using Gaussian noise filtered between 4 and 26 kHz (reproduced and modified, with permission, from Slee & Young, 2010).

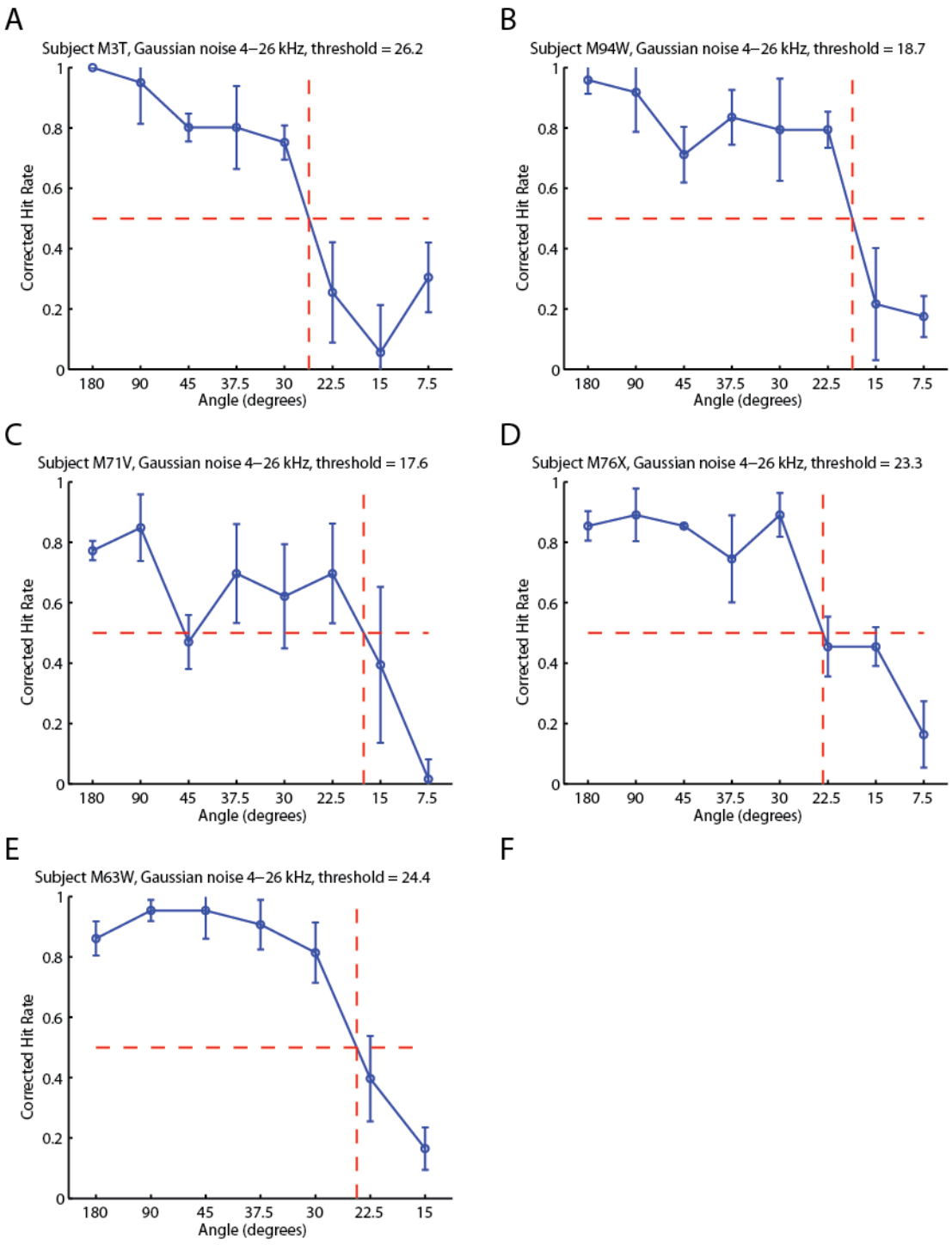


Figure 4.11. Frontal elevation discrimination: 4-26 kHz Gaussian noise. (**a-e**) Psychometric functions for marmosets detecting changes in frontal elevation location where the test stimulus was a Gaussian noise with no energy above the first spectral notch region.

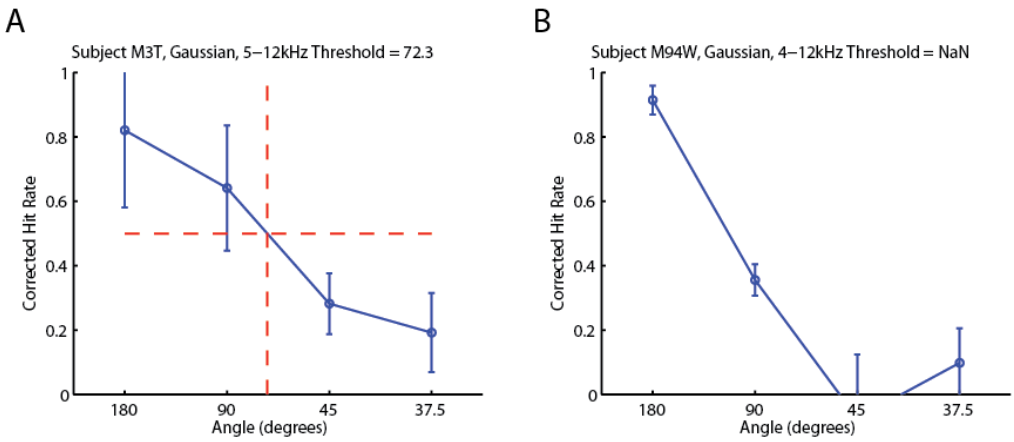
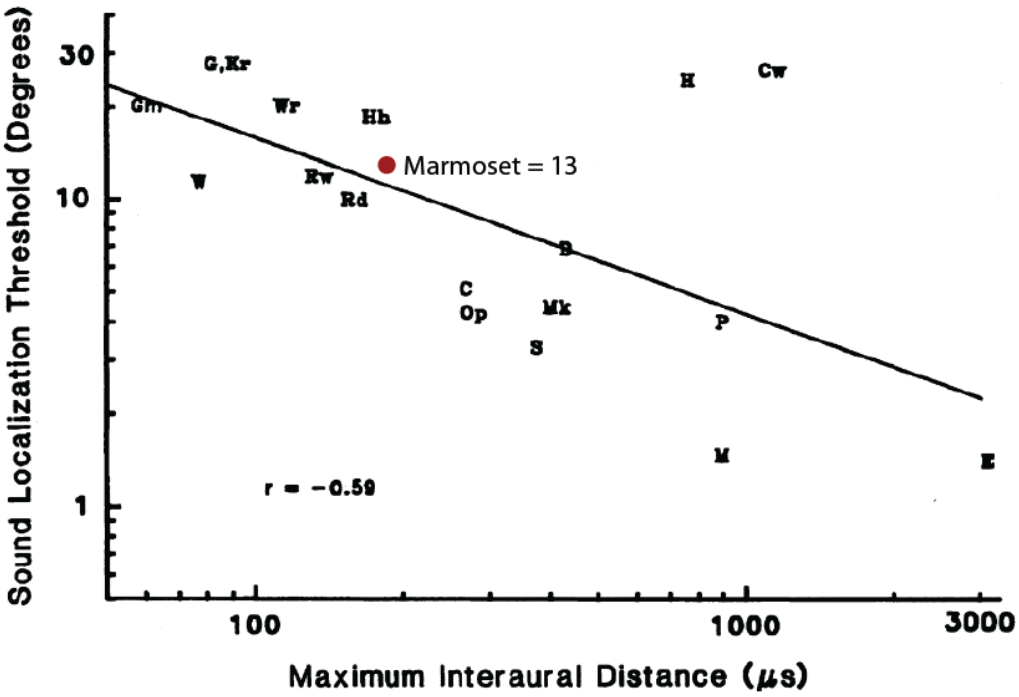


Figure 4.12. Frontal elevation discrimination: 4-12 kHz Gaussian noise. (**a,b**) Psychometric functions for marmosets detecting changes in frontal elevation location where the test stimulus was a Gaussian noise with no energy in or above the first spectral notch region.

A


Sound Localization Threshold (Degrees)
30
10
1
G,Kr
Gm
Wr
Hb
Marmoset = 13
W
Rw
Rd
H
Cw
D
C
Op
Mk
S
P
M
E
r = −0.59
100
1000
3000
Maximum Interaural Distance (μs)

B

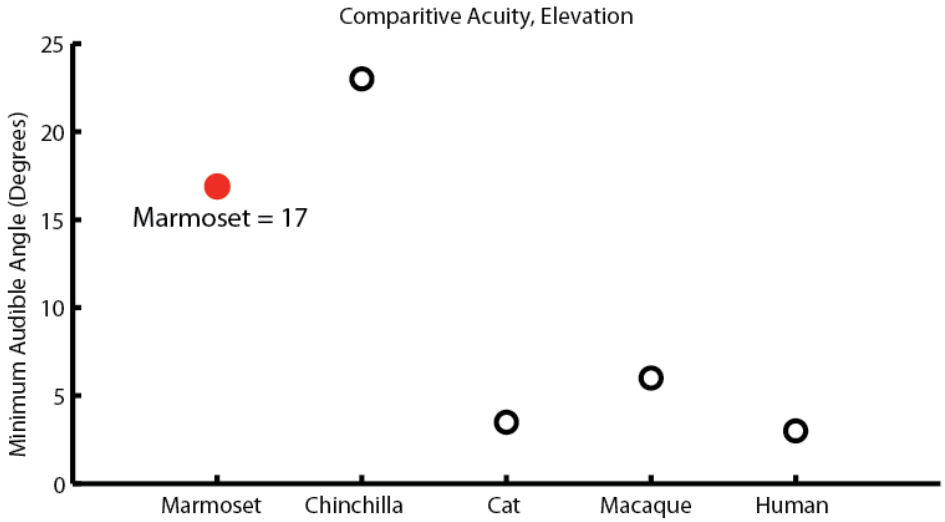

Comparitive Acuity, Elevation
Minimum Audible Angle (Degrees)
25
20
15
10
5
0
Marmoset = 17
Marmoset
Chinchilla
Cat
Macaque
Human

Figure 4.13. Comparative acuity. (**a**) Sound localization thresholds for broad-band stimuli as a function of head size in 18 mammals, plus the present value for marmosets. Marmosets are roughly in line with the trend gathered from previous studies. Gm, grasshopper mouse (Onychomys leucogaster); W, least weasel (Mustela nivalis); G, gerbil (Meriones unguiculatus); Kr, kangaroo rat (Dipodomys mer- riami); Rw, wild Norway rat (Rattus norvegicus); Rd, domestic Norway rat andWistar albino rat (R. norvegicus) Wr, wood rat (Neotoma floridiana); Hh, hedgehog (Paraechinus hypomelas); C, cat (Felis catus); Op, opossum (Didelphis virginiana); S, harbor seal (Phoca vitulina); Mk, rhesus and pig-tailed macaque monkeys (Ma- caca mulatta) and (M. nemestrina); D, dog (Canis canis); H, horse (Equus caballus); M, human (Homo sapiens) ; P, domestic pig (Sus scrofa); Cw, cattle (Bos taurus); E, elephant (Elephas maximus). Figure reproduced and modified, with permission, from Brown and May (2005). (**b**) Vertical sound localization discrimination thresholds in several mammals.

Chapter 5: Representation of sound location in a behaving primate

5.1 Results

5.1.1 Recording single-unit responses in the behaving condition

For this study, we recorded single unit activity while marmosets either listened passively to broadband sounds played from speakers on a complete sphere (see Chapter 3) or performed a sound location discrimination task (see Chapter 4) that included a subset of those locations. In the discrimination task, a trial was composed of an intertrial interval during which a variable number of sounds played from a background (reference) location, and a response interval in which sounds were alternated between one of four target locations and the background location. If a response (a lick at the feeding tube) was recorded during the response interval, the animal received a food reward. False alarms were measured with sham trials (sham trial probability was 30%). In most cases, sound level was roved in the behavior condition ±5dB or ±10dB to avoid the use of absolute level as a cue. For neurons which were highly sensitive to sound level, sound level was held constant in behavior. Four behavior target/background configurations, each with four targets and one background were used (see chapter 2); one of these configurations is illustrated in figure 5.1. Marmosets were head-fixed for all recordings. In analyses, firing rates in the behaving condition were compared to two passive conditions: one in which sounds are played in a random order from the full 24-speaker array (hereafter referred to as "passive"), and one in which sounds were played only from target/background locations in the same pattern as the behavior condition and with roved sound level (hereafter referred to as "control"). This second control condition was used to test for the potential confounding effects of stimulus order and sound level tuning (sound level was not roved in the passive condition). In the discrimination task, sounds were played

much more frequently from the background location than any of the target locations. Stimulus order effects can be highly variable, resulting in both suppression and facilitation (Ulanovsky et al. 2004; Bartlett & Wang 2005), and elevated firing rates for novel (or "surprising") stimuli have been frequently observed (Taaseh et al. 2011)**.** Unfortunately, stimulus order effects are not well studied for spatial location. All analyses were conducted on stimuli presented in successful discrimination ("hit") trials, unless otherwise noted.

5.1.2 Increased firing rates during behavior at target locations

The most common effect that was observed when comparing behaving and passive conditions was an increased firing rate in the behaving condition to one or more target locations. Figure 5.2.1-5 shows several example neurons which displayed significantly elevated firing rates in hit trials compared with the passive condition ($p < 0.05$) to several, but not all target locations when tested in various target/background configurations. Some units also displayed a significantly increased firing rate to the background location in one or more conditions. A similar effect was seen in the majority of units which had at least one significantly different firing rate between behavior (hit trials) and passive conditions. Figure 5.3A,B shows the population distribution of behaving vs. passive comparisons for each target and background location tested for each unit tested. Population analyses included neurons which displayed a driven firing rate ($p < .001$ and minimum mean rate of one spike per stimulus presentation) to at least one location in at least one condition. Although in most studies spontaneous rates have not been observed to change significantly in behavior (Miller et al. 1972; Benson & Hienz 1978; Otazu et al. 2009), several recent studies found increased spontaneous rates during behavior (Scott et al. 2007; Dong et al. 2013). In this sample population, there was a modest but significant increase in spontaneous firing rates when comparing hit trials to passive listening (mean modulation index = 0.2, $p < 0.001$). We therefore subtracted spontaneous firing rates from evoked rates for analyses. We also added an additional criterion that for a response to be considered significantly affected by

behavior, the neuron's firing rate to that location had to be "driven" in either the passive or behaving context. This was to avoid spurious statistical significance for very low firing rates by shifting spontaneous firing rates. All analyses in the following text treat each location tested in each neuron as a separate data point. There was a significant effect of behavior on firing rate to target locations, with a mean modulation index of 0.36 for all target/unit pairs (t-test, $p << 0.001$). Of 207 units tested, 70 had at least one significantly increased firing rate while 19 had at least one significantly decreased rate. For background locations (also in hit trials), both increases and decreases were observed, with a mean modulation index of 0.013 ($p = 0.9$).); 17 units had at least one significantly increased firing rate to backgrounds and 14 had at least one significantly decreased rate. The medians of modulation index values for significant data points for target and background locations were significantly different ($p < 0.001$), although in general direct comparisons of significant points between targets and backgrounds sets may be biased due to the fact that for the target locations, the smaller number of stimulus repetitions means that effects must be larger to become significant. These effects were consistent when comparing behaving and control conditions (figure 5.3C,D). In the control comparison, mean MI for targets was 0.50 ($p << 0.001$); of 111 units tested, 37 had at least one significantly increased firing rate while 2 had at least one significantly decreased rate. Mean MI for backgrounds was larger than for the passive comparison at 0.17 but did not reach significance ($p = 0.2$); 16 units had at least one significantly increased firing rate to backgrounds and 5 had at least one significantly decreased rate, and again medians of modulation index values for significant data points for target and background locations were significantly different ($p < 0.05$). As firing rates were averaged over trials comprised of multiple target/background alternations, and a correct discrimination prematurely ended a target trial, targets trials with correct responses will have fewer presentations. If there is strong facilitation of low-probability locations (i.e. "deviant" stimuli, (Taaseh et al. 2011)) independent of behavior state, this effect may be accentuated in the behaving condition. To test for this possibility, we performed a population analysis similar to

figure 5.3C**,** including both hits and misses ("targets" vs. control mean MI = 0.40, $p << .001$**,** figure 5.4). Modulation index for all data points with a significant effect in this condition was then recalculated using only the first target presentation within trials (mean MI = 0.38, $p << .001$). These results suggest that stimulus order effects were not the primary factor driving the observed increase in firing rates to target locations.

5.1.3 Distributed and additive nature of firing rate increases

In some previous studies, firing rate changes were correlated with underlying receptive fields (Benson & Hienz 1978; Lee & Middlebrooks 2011; Lee & Middlebrooks 2013), while in another effects appeared to be randomly distributed (Benson et al. 1981). It is possible that the interpretation of these effects may have benefited from a more complete sampling of spatial tuning. Here we looked where significant increases occurred along the passive response function (the passive firing rate scaled by the maximum passive firing rate) and found that firing rate increases could occur throughout spatial receptive fields, suggesting that effects were not limited to any specific receptive field area. Figure 5.5A plots firing rate increases at target locations in behavior vs. passive firing rate, both scaled by the maximum average firing rate in the passive condition. No significant correlation was found between the passive firing rate and the scaled rate increase ($r = 0.05$, $p = 0.6$), indicating that firing rates were increased in an additive rather than multiplicative manner. The same lack of rate increase vs. passive rate correlation existed for increases at background locations ($r = -0.01$, $p = 1$**,** figure 5.5B) and for comparisons with the control condition (targets: $r = 0.17$, $p = 0.2$; backgrounds: $r = 0.25$, $p = 0.3$, figure 5.5C,D). However, scaled firing rates for increased data points tended to be higher than that for the remaining data points (mean = 0.47 vs. 0.33, $p << 0.001$), suggesting that increases were more common when there was some level of response in the passive condition. In other words, it was uncommon for a neuron to fire strongly in the behaving condition if no response was observed in the passive condition.

Increases did not occur at all locations in units with increased responses, and this effect was not due to the predominance of non-driven locations which did not often show firing rate increases. Out of a minimum of 4 locations tested, the median number of significantly increased responses per significantly increased unit was 1 location, while the median number of driven locations in the same population of units was 3 ($p << 0.001$), indicating that increases did not occur throughout receptive fields; rather effects were location specific.

5.1.4 Target/background firing rate contrast

As the goal of the behavior task is to discriminate target and background locations, it seems reasonable to ask whether rate changes in behavior increase the firing rate to target locations relative to background locations. As firing rate increases were larger to target locations than to background locations, we hypothesized that this was the case. To test this hypothesis directly, the correlation between the scaled rate increase and the scaled contrast increase was calculated for both the passive comparison (figure 5.6A, $r = 0.82$, $p <<0 .001$) and control comparison (figure 5.6C, $r = 0.79$, $p << 0.001$). The high correlation between rate increases and contrast increases and the stark tendency of contrast increases to follow rate increases in general suggest that this hypothesis is correct. However, in cases where the target/background contrast was negative in the passive condition, it is possible that a positive change in contrast could actually be maladaptive for a hypothetical decoder attempting to use firing rates to detect a change in sound location**.** There was a small but significant positive correlation between passive contrast and contrast increase ($r = 0.34$, $p < 0.001$). Figure 5.6B contains a plot of scaled contrast increase vs. scaled passive contrast. This were also consistent when comparing behaving firing rates with the control condition ($r = 0.57$, $p < 0.001$; Figure 5.6D). These results are in contrast with the present finding that firing rate increases could occur throughout the receptive field, but suggest that firing rate increases at the population level act to increase target/background contrast in neurons which already may provide a positive detection signal for target locations.

5.1.5 Comparison of behavioral modulation of neural responses between areas A1, R/RT, and CM/CL

Several studies have shown that the distributions of spatial tuning properties vary quantitatively (but not qualitatively) between auditory areas along the rostral-caudal axis, with neurons in caudal areas, statistically, displaying higher selectivity for spatial locations than those in rostral areas and primary auditory cortex (Rauschecker & Tian 2000; Stecker et al. 2003; Woods et al. 2006). However, only one study has compared the effects of engagement in a spatial auditory task between these areas (Lee & Middlebrooks 2013). Units were separated by area using frequency map gradient reversal points (see methods), and a significant effect of area on modulation index was observed when comparing behaving and passive conditions (ANOVA, $p < 0.001$; $p < 0.01$ for all two-area comparisons), finding the largest values in CM/CL, and the smallest in R/RT (figure 5.7A,B). A striking result obtained by separating the data by area was that the effects of behavior on targets seemed to be nonexistent in areas R/RT (mean MI = −0.04, $p = 0.8$). When this analysis was performed using behavior/control comparisons, the trend was similar, however differences were no longer significant except for the CM/CL to R/RT comparison (ANOVA, $p = 0.09$; CM/CL vs. R/RT, $p < 0.05$ figure 5.7C,D). This lack of effect is likely partially due to the reduced statistical power of the target/control comparisons (n = 37 units compared with 86 for the target/passive comparison), but also due to the increased effect in areas R/RT. The mean modulation index for the behavior/control comparison in R/RT was 0.34 and yielded a significant effect, (t-test, $p < 0.05$), though the comparison between behavior/control and behavior/passive was not significant ($p = 0.16$). Comparing plots of firing rates for hits vs. passive and hits vs. control (figure 5.3A,C), the most obvious difference is the lack of significantly decreased rates in the behavior vs. control condition for all areas, suggesting that effects of stimulus order may play a role in depressing firing rates in the behaving and control condition, particularly in areas R/RT. Finally, the analysis of behavior effects on target/background contrast was repeated for individual

auditory areas. In CM/CL, there was a positive correlation between firing rate increase and passive firing rate ($r = 0.35$, $p < .05$), suggesting that behavior effects in this area were both additive and multiplicative (figure 5.8A). This trend was stronger when comparing behaving and control conditions ($r = .64$, $p < .05$, figure 5.8B). This result indicates that there may be a qualitative difference in the way that area CM/CL processes spatial information in behavior.

5.1.6 Relationship between tuning properties and behavior effects

Correlations between behavioral sensitivity and physiological response properties are not well established. Two spatial studies have observed distinctions between behaviorally sensitive and non-sensitive neurons based on evoked firing rates. First, in macaques discriminating left vs. right ear stimuli, evoked responses increased only for contralateral preferring neurons (Benson & Hienz 1978). More recently, receptive fields were shown to narrow in a spatial discrimination task primarily in neurons with broad spatial receptive fields (Lee & Middlebrooks 2011). We asked whether any connections could be found between tuning properties and behavioral sensitivity. These tests were done by dividing units into two or more subpopulations, then comparing the distribution of affects based on the modulation index.

First, we tested for an effect of laterality preference; classifying neurons by lateral centroid (see methods). Lateral centroids (which could range from −90° to 90° see methods) of less than −22.5° were considered contralateral, greater than 22.5 were ipsilateral, and the remaining were considered central. There was no significant effect of laterality preference on modulation index (ANOVA, $p = 0.7$), and proportions of units with significant effects were similar: 24 of 76 contralateral units, 40 of 108 central, and 6 of 24 ipsilateral units had significant increases. Ipsilateral units were much less well driven in this sample population, making comparisons difficult, but central and contralateral preferring neurons seemed to respond similarly in behavior (figure 5.9A,B). The second grouping was with respect to SRF tuning area.

Based on the distribution of tuning area measured in the passive condition (see chapter 3), neurons with a tuning area less than 0.3 were considered "narrow," and those with tuning area greater than 0.3 "broad." The effect size for narrowly tuned neurons was slightly larger than for broadly tuned neurons (mean MI = 0.45 for narrow and 0.26 for broad; $p < .05$), but the fraction of narrowly tuned neurons with significant increases was smaller (38/131 for narrow vs. 32/78 for broad). The smaller fraction of narrowly tuned neurons with significant increases may be due to the fact that it was uncommon for a neuron to fire strongly in the behaving condition if no response was observed in the passive condition, and narrowly tuned neurons by definition responded to fewer locations. Also, as R/RT had a higher proportion of broadly tuned neurons, and R/RT had a small behavior effect size, this comparison was repeated on units in A1 and CM/CL only**.** In A1 and CM/CL, The effects for narrowly and broadly tuned neurons were more similar (mean MI = 0.47 for narrow and 0.39 for broad, $p = 0.37$; figure 5.9C,D). The final comparison was between monotonic and nonmonotonic neurons. Here the threshold for a neuron to be considered "monotonic" was a monotonicity index of 0.5 (see see chapter 6 for a discussion on monotonicity observed in this study). There was a tendency for nonmonotonic neurons to have mixed effects when comparing the behaving to the passive condition, although there was not a significant difference between the two populations (mean MI = 0.17 for nonmonotonic and 0.36 for monotonic; $p = 0.13$). We again suspected that this effect may have been due to the predominance of nonmonotonic neurons in R/RT, so the test was repeated using neurons in A1 and CM/CL only. Here there was no significant difference in MI (mean MI = 0.36 for nonmonotonic and 0.42 for monotonic; $p = 0.6$), but the fraction of significantly increased units was smaller for nonmonotonic neurons (11/42 vs. 18/37 figure 5.10A,B). This may be due to the tendency of nonmonotonic neurons to have narrower tuning areas, (see chapter 3), which, as mentioned previously, can lead to fewer driven locations.

5.1.7 Relationship between task performance and behavior effects

Behavior performance was variable between locations and behavior sessions; however the present study was not designed to determine whether responses to target locations in auditory cortex were predictive of behavior responses for target trials. Behavior performance was found to be a poor predictor of the effects of behavior on firing rates. In this analyses, data from hit and miss trials was included. Figure 5.11A shows modulation index vs. hit rate for all data points from neurons with at least one significant change (including nonsignificant data points). The correlation value was small and nonsignificant ($r = 0.075$, $p = 0.1$). Figure 5.11B shows modulation index vs. mean session hit rate for all data points from all neurons. Although statistically significant, the correlation was equally small ($r = 0.075$, $p < 0.01$). Behavior sessions in which animals were not engaged in the task due to apparent lack of motivation were rejected.

There were two notable observations for background stimuli. First, we compared firing rates at background locations in error trials (a lick outside of a target trial) and non-error trials (ITI periods and sham trials in which no lick was recorded), and found virtually no difference across the population, although the effect was statistically significant (mean MI = 0.064, $p < 0.05$; figure 5.12A); However, comparing spontaneous firing rate in the same two conditions, a significant effect was observed (mean MI = 0.37, $p << 0.001$, figure 5.12B). Together, these results suggest that baseline activity was generally increased in some neurons prior to an incorrect response, evoked (spontaneous-rate subtracted) responses were not. There was no significant effect comparing spontaneous firing rate in hit trials vs. miss target trials (mean MI = 0.17, $p = 0.08$, figure 5.12C).

5.1.8 Behavioral modulation at vertical vs. horizontal discrimination locations

Of the 4 target locations used in the discrimination task, 2 locations differed from the background only in azimuth, while 2 others differed in azimuth and elevation. It is possible that the effects of

behavior could be different for these two sets of targets. Here again data from both hit and miss trials are included to control differing performance levels between the two groups (mean hit rate was 0.53 for vertical targets and 0.70 for lateral targets; $p << 0.001$). However, there was not a large difference in the number of significantly increased responses (71 lateral, 80 vertical) or modulation index for significantly modulated data points (mean MI lateral = 0.32, mean MI vertical = 0.26, $p = 0.33$) between the two groups, indicating that the effects were not substantially different.

5.2 Discussion

5.2.1 Summary of findings

We recorded single-unit responses in auditory cortex of marmosets while they performed a spatial discrimination task in different regions of the full spatial field. Comparing these responses to those measured while marmosets sat passively, a subset of neurons was observed to display increased firing rates to one or more target locations during task engagement. Effects at background locations were mixed. As the behavior task involved a specific stimulus order which was different than that used to measure spatial receptive fields, we measured responses in an additional passive control condition in which stimulus order was identical to that of the behavior condition. Effects were similar, even slightly larger, when comparing rates with the control condition, indicating that observed effects were not driven by stimulus order effects in most neurons. Increases occurred both within and outside of the classical spatial receptive field (typically defined as half-maximal firing rate area), but were less common when there was no response to a location in the passive condition. Firing increases served to increase firing rates relative to the reference location. Comparing effects of behavior between rostral (R/RT), caudal (CM/CL), and primary (A1) auditory areas, the largest effects were observed in CM/CL and the smallest in R/RT.

5.2.2 Location specificity of behavioral modulation

The observation that firing rate increases occurred at some and not all locations and were also distributed throughout receptive fields indicates that effects in this task were largely location (or stimulus) specific. This has been observed previously in one study of spatial behavior (Benson et al. 1981), in which subjects were required to localize a sound source. In others, effects tended to occur throughout (Scott et al. 2007) or either in preferred or non-preferred portions of receptive fields (Lee & Middlebrooks 2011; Benson & Hienz 1978), indicating neuron-specific effects.

Several studies in behaving ferrets have shown consistent stimulus-specific effects on frequency tuning, but not neuron-specific effects (Fritz et al. 2003; Fritz et al. 2007). This dichotomy has also been observed in studies of spatial attention in the visual system, with neuron-specific effects including, but not limited to, multiplicative gain (Treue & Martínez Trujillo 1999) and receptive field sharpening (Spitzer et al. 1988), and stimulus-specific effects exemplified by receptive field shifts (Womelsdorf et al. 2006).

The stimulus-specific effects raise the interesting question of what mechanisms underlie the changes observed here. For instance, a simple reduction in inhibitory strength would not explain the effects. Although it is possible that these observations can be explained by neuron-specific effects acting at some input of the neurons recorded in this study, the spatially restricted effects would seem to be the result of modulation of input neurons with small receptive fields. However, receptive fields measured in subcortical areas are not narrower than those in auditory cortex (Aitkin & Jones 1992; Delgutte & Joris 1999). An intriguing possibility is domain specific modulation of synaptic integration. It is now well known that dendritic integration is governed by active electrical properties in pyramidal cells of the cortex, and recent work has found that these active domains can be compartmentalized by additional active elements (Xu et al. 2012; Harnett et al. 2013). Furthermore, it has been shown that neuromodulation can be spatially restricted to specific dendritic compartments (Hasselmo & Schnell 1994; Nakamura et al. 1999). Clearly, further studies are necessary to understand the contrasting mechanisms of neuron-specific and stimulus-specific modulation across multiple sensory modalities.

5.2.3 Increases at target locations vs. decreases at background locations

Another dichotomy in observations of the effects of behavioral engagement on firing rates in auditory cortex is that some studies have observed increased responses, generally at target locations (Miller et al. 1972; Benson & Hienz 1978; Benson et al. 1981; Dong et al. 2013), while

more recently several studies have shown decreases, primarily to background locations (Otazu et al. 2009; Lee & Middlebrooks 2011). Although stark differences have been observed due to task structure, such as between appetitive and aversive tasks (David et al. 2012), these differences were all in studies using positive reinforcement. One study was quite similar to ours in structure yet had quite dissimilar results: in cats performing a sound elevation discrimination task, responses were often decreased at non-preferred locations, while effects at preferred locations were mixed (Lee & Middlebrooks 2011). It is possible that the "preferred" locations in this study were not truly preferred, as receptive fields were only sampled along the azimuth dimension at zero elevation. The results discussed in chapter 3 suggest that a number of neurons have preferred locations off of the horizon. Still, all background stimuli in this study were located on at zero elevation, and a consistent decrease in firing rates to background locations was not observed. This also suggests that the difference is not primarily due to the analytical focus on targets vs. backgrounds. We can think of at two other possibilities. First, it may be the case that behavior effects are highly dependent on very specific details of task structure, such as the distributed backgrounds in the cat study vs. the distributed targets in ours. Alternately, the neural populations targeted could be different. In the cat study, most neurons observed in the passive condition responded broadly with onset responses. However, our neural population also included onset responders and broadly tuned neurons (at least to the test stimuli), and these neurons could also show increased firing rates (for example, see figure 5.2.2 and figure 5.9C). Still, the sharp difference in the tuning distribution between the two studies could be indicative of a bias in recording methodology between the two studies.

5.2.4 Increased firing rates at target locations relative to backgrounds

While firing rate increases could occur throughout receptive fields, the tendency for firing rate changes to occur more consistently and strongly at target locations rather than at both target and background locations suggests that these effects acted to tailor spatial representation for the

specific behavior task. Further supporting this notion is the fact that increases were more likely to be larger where target/background contrast was positive in the passive condition. Thus, both neurons' spatial tuning and the arrangement of target and background locations influenced the firing changes in behavior. Previous studies of mammals performing sound location tasks have observed similar seemingly optimizing changes in spatial tuning. In the first, neural responses were recorded in macaques performing a dichotic listening task in which they were instructed to respond sounds played to the left or right ear. In this task a population of contralateral preferring neurons responded more strongly to contralateral locations when the target was contralateral (Benson & Hienz 1978). More recently, cats trained to discriminate sound elevation along the complete azimuthal dimension displayed depressed responses at non-preferred locations (Lee & Middlebrooks 2011). In both of these cases changes occurred which were optimized to the task in the context of the underlying spatial (or binaural) tuning. It has been suggested that this is an underlying high level behavior of spatial processing in auditory cortex (Lee & Middlebrooks 2011); our results seem to add support to this hypothesis.

5.2.5 Comparison of effects between areas A1, R/RT, and CM/CL

Several studies have shown that the distributions of spatial tuning properties vary quantitatively (but not qualitatively) between auditory areas along the rostral-caudal axis, with neurons in caudal areas, statistically, displaying higher selectivity for spatial locations than those in rostral areas and primary auditory cortex (Rauschecker & Tian 2000; Stecker et al. 2003; Woods et al. 2006; Lee & Middlebrooks 2013). If, as hypothesized, caudal areas are especially important in sound localization behaviors, one would expect meaningful changes in the way tasks involving sound location affect caudal vs. primary and rostral auditory cortex. A recent study comparing effects between A1, PAF, and DZ in cats found a lower fraction of neurons in caudal areas PAF and DZ which sharpened tuning compared with A1 (Lee & Middlebrooks 2013). This report represents the first time behavior effects on spatial responses in primates have been directly

compared between multiple auditory areas including significant recordings from the caudal area, although effects of behavior outside of primary auditory cortex have been measured previously, primarily in rostral and lateral areas (Benson et al. 1981; Benson & Hienz 1978). For the most part, quantitative, not qualitative difference in effects between areas mirrors quantitative, but not qualitative differences in tuning properties between areas (Stecker et al. 2003; Stecker et al. 2005; Woods et al. 2006; Zhou & Wang 2012). Although the data do suggest the possibility of multiplicative in addition to additive gain in areas CM/CL, this effect was small. The most dramatic difference in effect size was the apparent lack of consistent behavior effects in R/RT; however it is possible that these differences were confounded by stimulus order effects. When behaving firing rates were compared with the control condition with identical stimulus order, increases were apparent, albeit still smaller than in A1 and CM/CL. It has previously been shown that neurons in rostral areas integrate stimulus information over longer time windows than units in primary auditory cortex (Bendor & Wang 2007; Bendor & Wang 2008), but these studies focused on rapidly time-varying stimulus patterns rather than interactions between successively presented sounds on the order of hundreds of milliseconds to seconds. To our knowledge, no study has specifically compared the effects of stimulus interaction over long time scales across the rostral-caudal axis.

5.2.6 Correlation of behavior and single neuron activity

Neural stimulation experiments have established a causal link between activity in sensory cortex and behavior (Yang et al. 2008; Jazayeri et al. 2012; Znamenskiy & Zador 2013), even in single neurons (Houweling & Brecht 2008). Several studies have also shown that behavioral responses can be predicted to varying degrees using single neuron responses (Britten & Newsome 1996; Dodd et al. 2001; Nienborg & Cumming 2006). Due to the use of multiple target conditions and time constraints in recording from a given neuron, the present study was not well-suited to determine whether responses to target locations in auditory cortex were predictive of behavior

responses for target trials. However, while the observation that behavioral effects were not well predicted by behavioral performance does not rule out that such a correlation may exist, it does not support it. The observations regarding spontaneous firing rates, however, are curious. It has previously been suggested that neurons in auditory cortex can respond to non-auditory events, including behavior responses (Brosch et al. 2005). It is unlikely that the effects here have a similar cause. First, previously observed non-auditory responses were time-locked to specific events. False alarm trials in this study were averaged over several successive stimuli, so changes on very short time scales would likely be averaged out. Also, there was not a comparable increase in spontaneous activity in hit trials vs. miss trials, although the differences in behavioral responses were the same. It is possible that this increase in baseline firing has some causal role in perceptual errors leading to erroneous behavioral responses.

5.2.7 Task design limitations

We can think of two limitations of the task used in the current study. First, the task was a location discrimination task rather than a localization task. In this case, it is possible that the task was not a spatial task from the animals' perspective. We took several precautions to reduce the likelihood that subjects were making discriminations based on sound qualities which were not perceived as spatial differences. First, sound level was roved (10 or 20dB) in most cases to reduce the use of absolute level cues. Second, when the target sound was noise, unique samples were generated for every stimulus presentation to reduce the reliance on subtle changes in sound spectrum with changing location. Finally, speakers were chosen which were closely matched spectrally to further reduce the number of non-HRTF related spectral cues. Still, animals have been shown to be able to use HRTF cues which are not useful for absolute localization in a location discrimination task (Huang & May 1996b; Huang & May 1996a), so the possibility that subject's were making a sound quality discrimination cannot be ruled out. However, the tendency for firing rates to increase at somewhat arbitrary locations within receptive fields has been observed

previously, specifically when comparing responses during an absolute localization task and a non-spatial task (Benson et al. 1981).

A second limitation of the current study is that firing rate comparisons were made between awake and passive conditions rather than between different behaving conditions. Therefore, it is not known whether the observed effects were a result of engagement in the location discrimination task specifically as opposed to general alertness. However, the finding that firing rates to target locations increased relative to background locations argues that the changes were an effect of engagement in this specific task rather than a general effect of arousal. Additionally, it may be difficult to interpret differences in neural responses between different tasks because of other factors besides modality that may differ between tasks. Specifically, task difficulty is known to modulate the effects of behavior on neural response properties (Spitzer et al. 1988; Spitzer & Richmond 1991; Atiani et al. 2009). Still, it is possible that the effects observed may be an effect of general arousal, as previous studies have shown that effects which happen to appear to be optimal for a particular task may also be observed to a lesser extent in other tasks when compared with passive listening (Lee & Middlebrooks 2011). Therefore, we believe it is likely that the effects observed in this report are a combination of general arousal as well as engagement in the specific task performed.

5.2.8 Conclusions

The varied but significant effects of behavioral context on spatial receptive fields in auditory cortex suggest that spatial tuning in the passive condition provides only part of the picture of spatial representation in auditory cortex. In fact, these and other behavior studies may reconcile disparate observations regarding spatial tuning in auditory cortex. For example, studies of spatial processing in anesthetized animals have observed a preponderance of neurons with very broad spatial receptive fields which almost universally increase in size with increasing sound level,

leading some to hypothesize that spatial decoding assumes a "panoramic" encoding scheme (Middlebrooks et al. 1994; Stecker et al. 2005). However, in awake animals, spatial receptive fields tend to be smaller and do not uniformly increase in size with sound level (Mickey & Middlebrooks 2003; Zhou & Wang 2012). The observation that large firing rate increases can occur far from best location may be evidence of broader inputs that are masked in the passive awake condition when compared to the anesthetized state. Conversely, another behavior task may lead to still more selective tuning compared to the passive state (Lee & Middlebrooks 2011). We therefore believe that the true nature of spatial representation may not be well understood by studying non-behaving subjects, and that a complete picture will require further behavior studies.

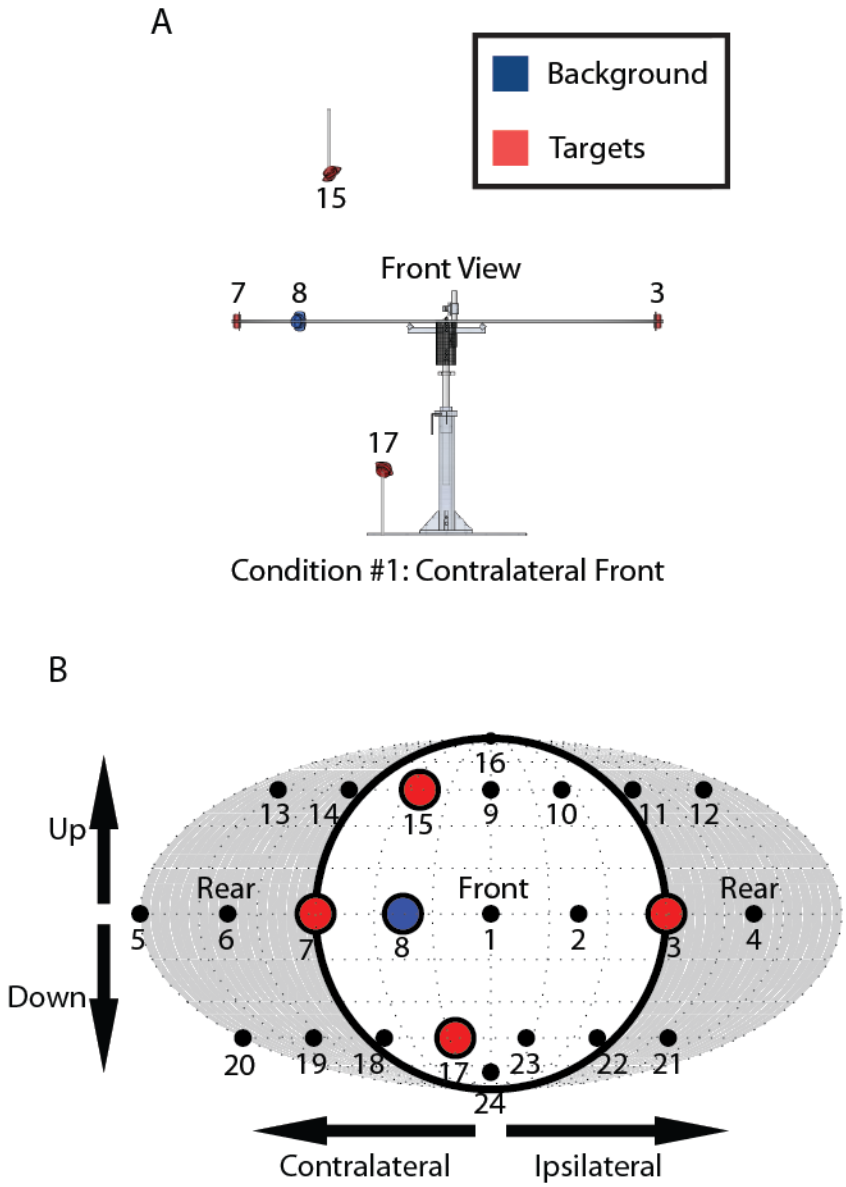


Figure 5.1. Task layout. (**a**) Speaker layout for one of the four behavior conditions The background location was 45° off of midline and the target locations were the most lateral positions (±90°; same in all conditions), and also 45° above and below the horizon, but in the same azimuthal quadrant as the background location. For targets above and below the background location, their azimuth locations were either 51° or 25.5° off of the midline (one of each per condition). (**b**) Fournier map projection of complete 24-speaker spatial array with speakers from (**a**) highlighted.

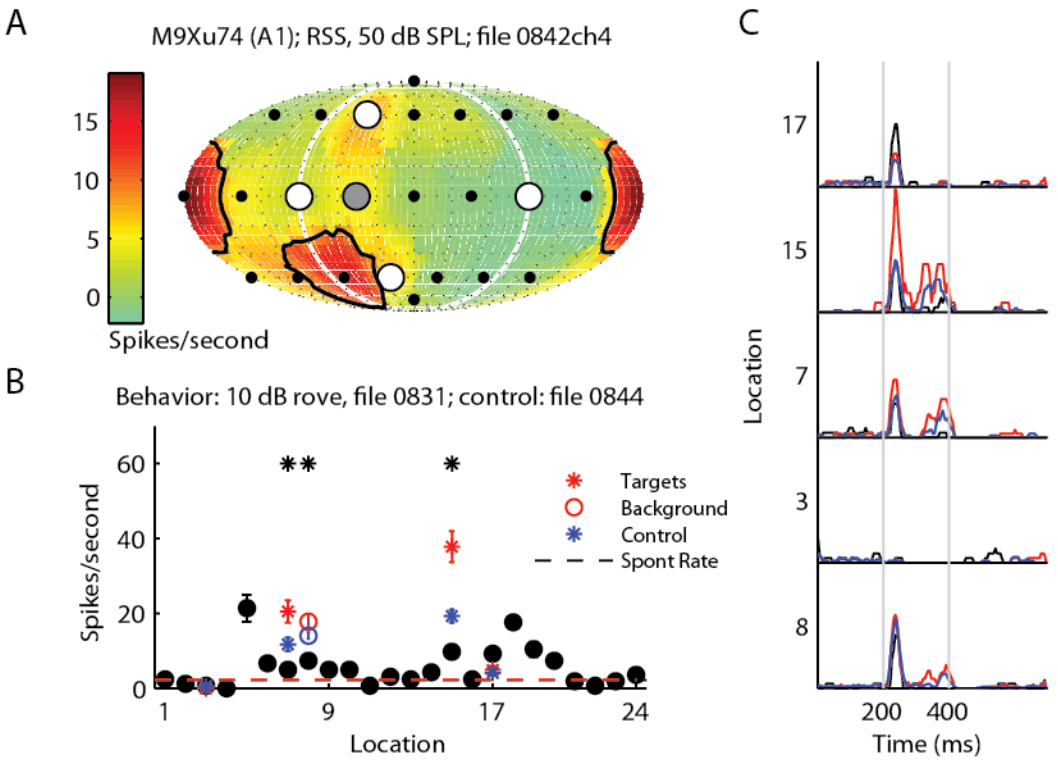


Figure 5.2.1. Example neurons displaying behavior modulation (1). Firing rates increase to some, but not all locations during spatial discrimination. (**a**) Passive spatial receptive field. This unit has a complex spatial receptive field, with responses distributed among contralateral and rear locations. Speaker locations, including target and background locations of the discrimination task configuration in figure 1 are overlaid on the response map.(**b**) Comparison of firing rates measured in the passive (black circles) active behaving (red stars and circle) and control (blue stars and circle) conditions. Black asterisks indicate locations with significant differences in firing rate between the behaving and both the passive and control conditions. Spontaneous firing rates in behaving and passive conditions are indicated with colored dashed lines; this neuron did not display a change in spontaneous firing rate in behavior. This neuron displayed increased firing rates to two of four targets, as well as the background location, in the behaving condition. Error bars denote standard error of mean. (**c**) Peristimulus time histogram for target and background locations in each condition.

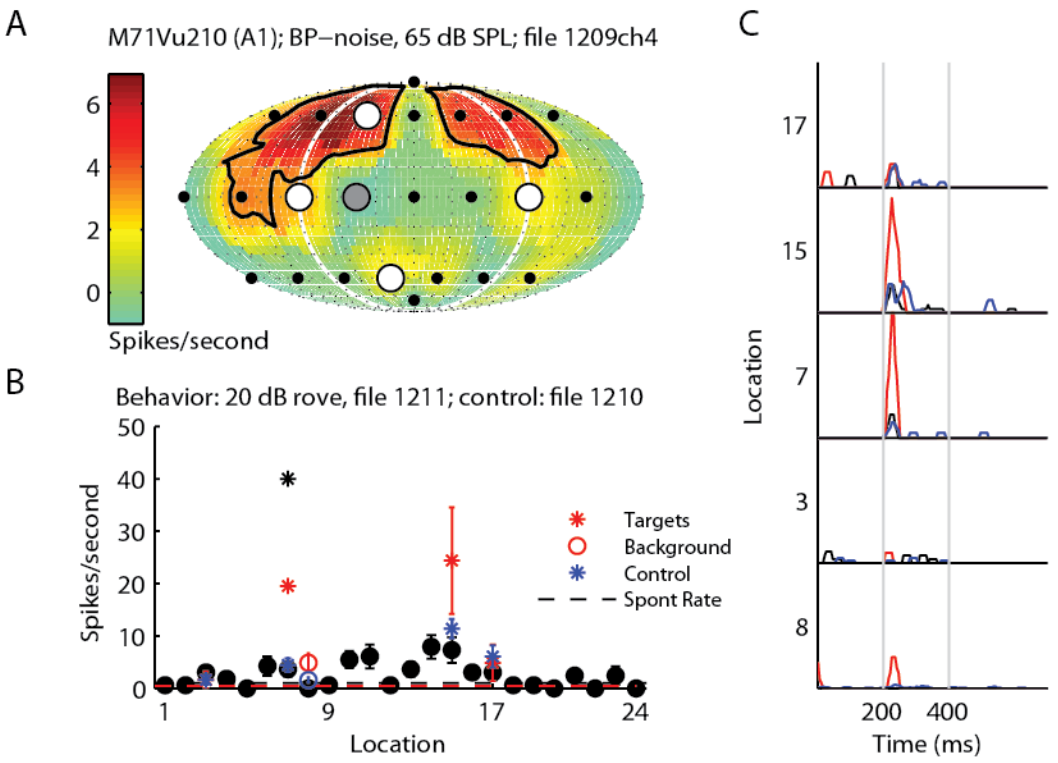


Figure 5.2.2. Example neurons displaying behavior modulation (2). Second example unit showing increased firing rates at a subset of locations in the behavior condition. (**a**) Passive spatial receptive field. This unit has a complex spatial receptive field, with responses distributed among contralateral and ipsilateral locations. Speaker locations, including target and background locations of the discrimination task configuration in figure 1 are overlaid on the response map.(**b**) Comparison of firing rates measured in the passive (black circles) active behaving (red stars and circle) and control (blue stars and circle) conditions. Black asterisks indicate locations with significant differences in firing rate between the behaving and both the passive and control conditions. Spontaneous firing rates in behaving and passive conditions are indicated with colored dashed lines; this neuron did not display a change in spontaneous firing rate in behavior. This neuron displayed increased firing rates significantly to one of four targets in the behaving condition. Error bars denote standard error of mean. (**c**) Peristimulus time histogram for target and background locations in each condition.

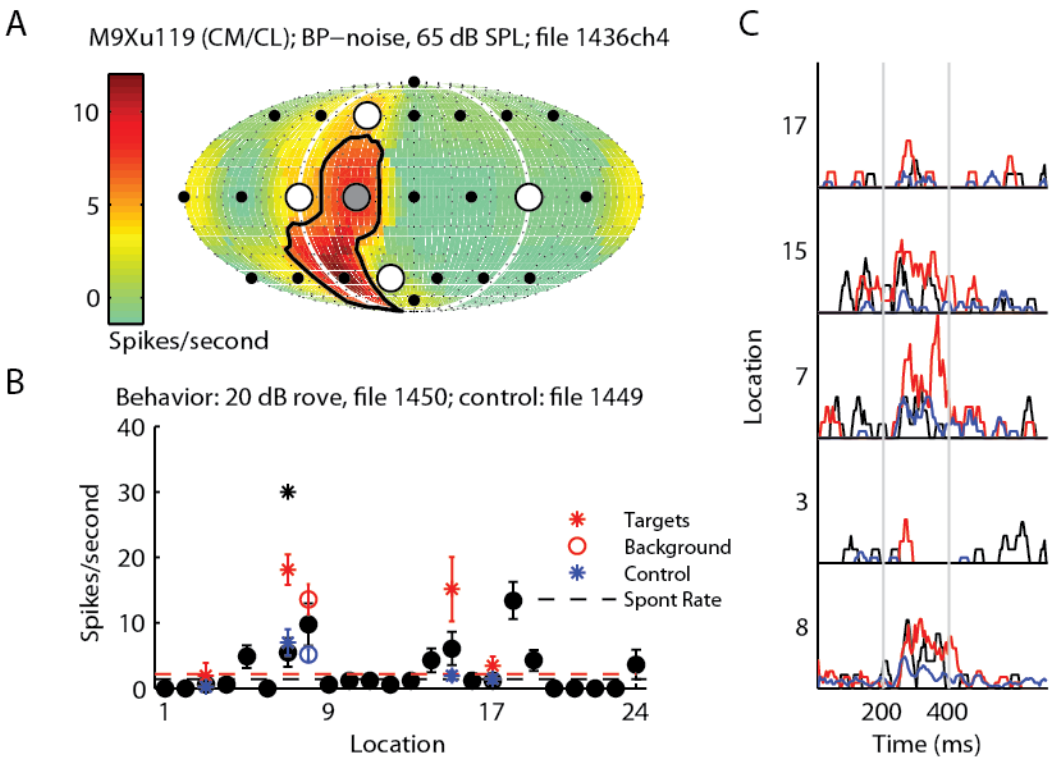


Figure 5.2.3. Example neurons displaying behavior modulation (3). Third example unit showing increased firing rates at a subset of locations in the behavior condition. (**a**) Passive spatial receptive field. This unit preferred frontal contralateral locations. Speaker locations, including target and background locations of the discrimination task configuration in figure 1 are overlaid on the response map.(**b**) Comparison of firing rates measured in the passive (black circles) active behaving (red stars and circle) and control (blue stars and circle) conditions. Black asterisks indicate locations with significant differences in firing rate between the behaving and both the passive and control conditions. Spontaneous firing rates in behaving and passive conditions are indicated with colored dashed lines; this neuron did not display a change in spontaneous firing rate in behavior. This neuron displayed increased firing rates significantly to one of four targets in the behaving condition. Error bars denote standard error of mean. (**c**) Peristimulus time histogram for target and background locations in each condition.

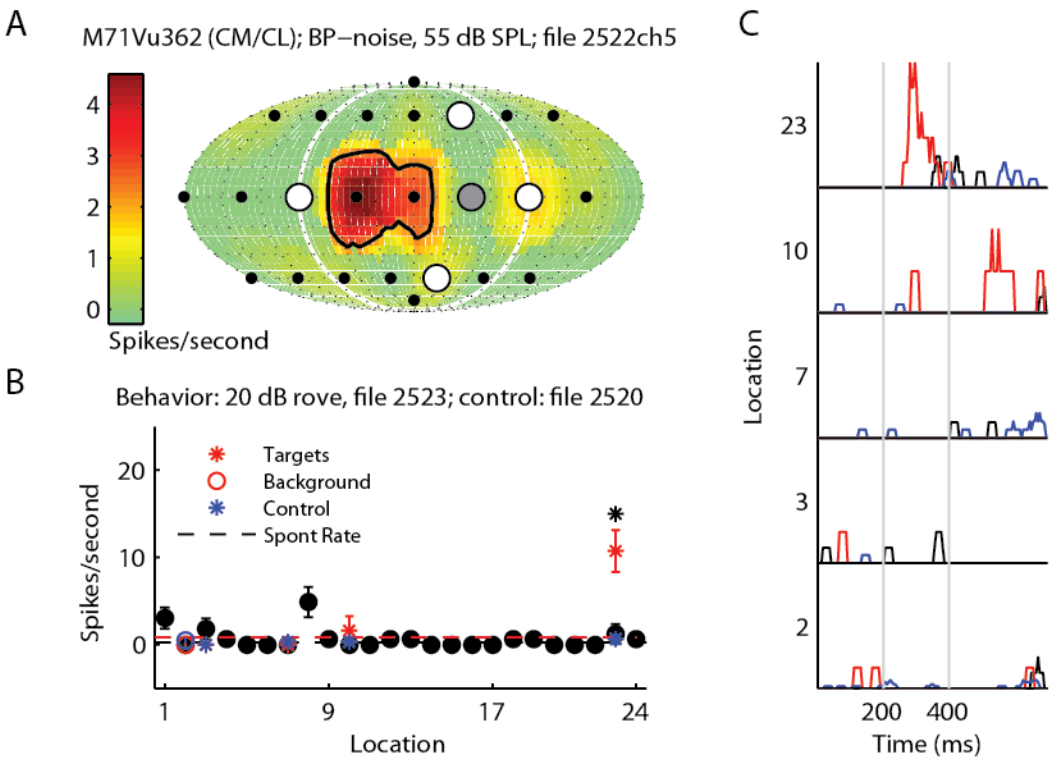


Figure 5.2.4. Example neurons displaying behavior modulation (4). An example unit showing increased firing rates at a contralateral and non-preferred location. (**a**) Passive spatial receptive field. This unit preferred frontal contralateral locations. Speaker locations, including target and background locations of the discrimination task configuration in figure 1 are overlaid on the response map.(**b**) Comparison of firing rates measured in the passive (black circles) active behaving (red stars and circle) and control (blue stars and circle) conditions. Black asterisks indicate locations with significant differences in firing rate between the behaving and both the passive and control conditions. Spontaneous firing rates in behaving and passive conditions are indicated with colored dashed lines; this neuron did not display a change in spontaneous firing rate in behavior. This neuron displayed increased firing rates significantly to one of four targets in the behaving condition. This neuron was not well driven in the passive condition. Error bars denote standard error of mean. (**c**) Peristimulus time histogram for target and background locations in each condition.

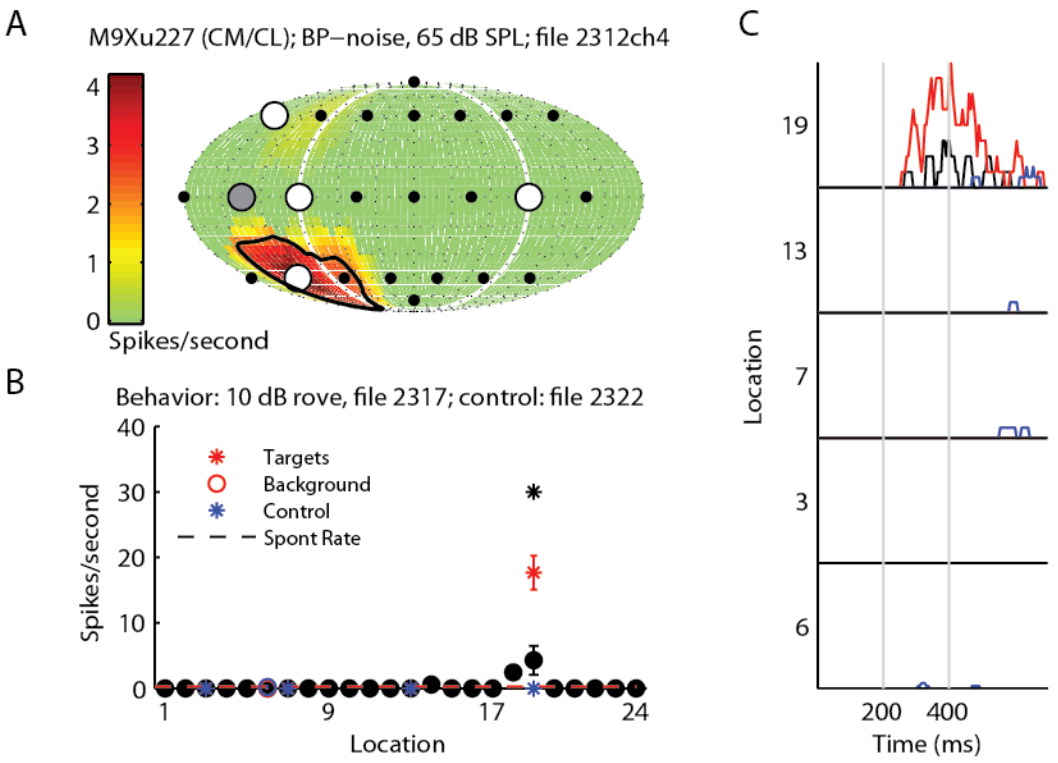


Figure 5.2.5. Example neurons displaying behavior modulation (5). An example unit showing increased firing rates at a rear location. (**a**) Passive spatial receptive field. This unit preferred frontal contralateral locations. Speaker locations, including target and background locations of the discrimination task configuration in figure 1 are overlaid on the response map.(**b**) Comparison of firing rates measured in the passive (black circles) active behaving (red stars and circle) and control (blue stars and circle) conditions. Black asterisks indicate locations with significant differences in firing rate between the behaving and both the passive and control conditions. Spontaneous firing rates in behaving and passive conditions are indicated with colored dashed lines; this neuron did not display a change in spontaneous firing rate in behavior. This neuron displayed increased firing rates significantly to one of four targets in the behaving condition. This neuron was also not well driven in the passive condition. Error bars denote standard error of mean. (**c**) Peristimulus time histogram for target and background locations in each condition.

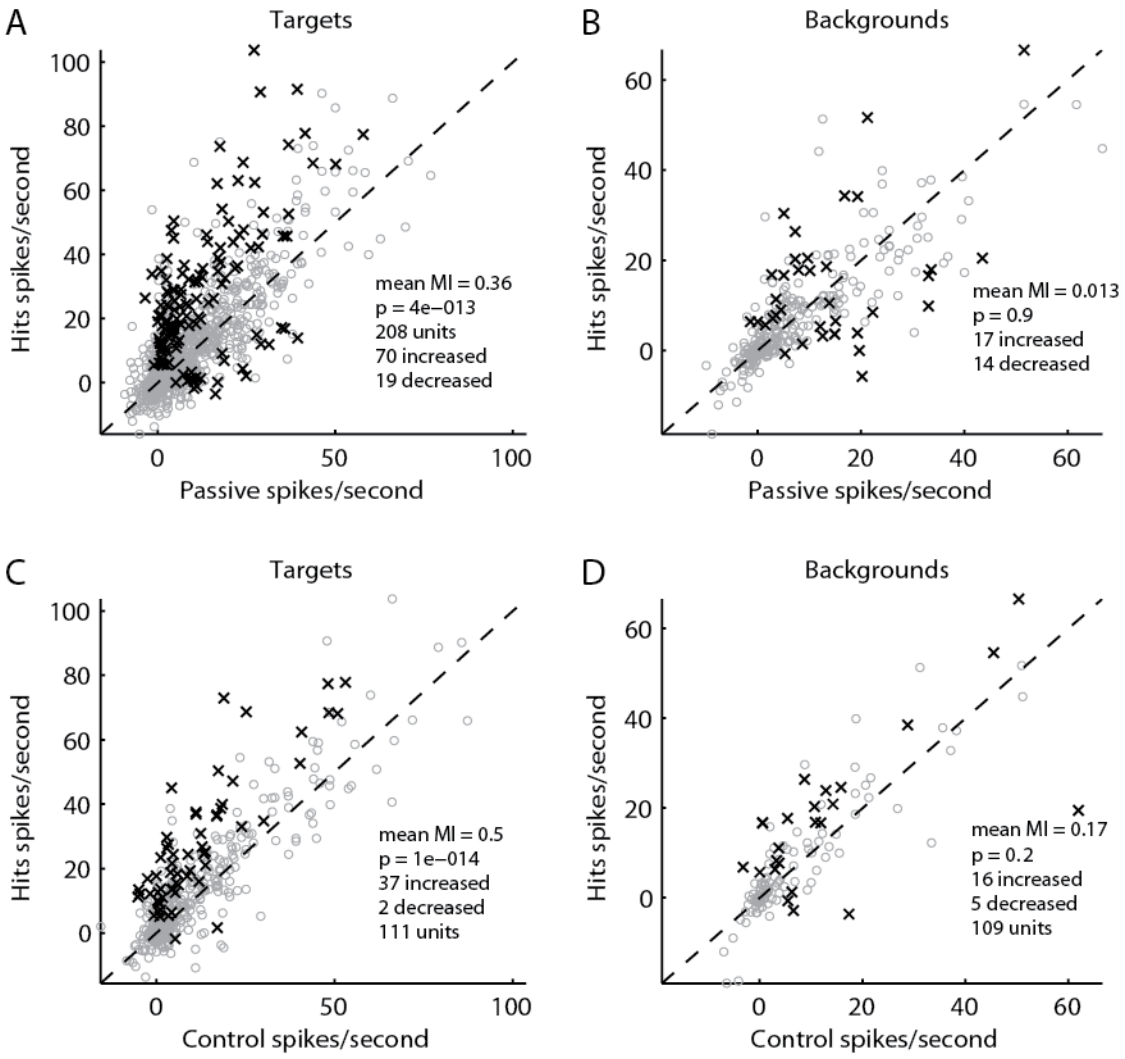


Figure 5.3. Population analysis: modulation index. Comparisons of firing rates in behaving and passive conditions for all units at all locations tested. (**a**) Target locations. In the majority of neurons displaying behavior modulation, firing rates were elevated in the behaving condition compared with the passive condition. Black x's represent significant data points ($p < 0.05$). All firing rates are spontaneous rate subtracted. (**b**) Background locations. Effects on background locations were mixed, with no clear trend to increase or decrease firing rates in the behaving condition. (**c,d**) Comparison of firing rates in behavior (hit trials) vs. a control condition which was identical to the to the behavior condition except that stimuli did not stop if the animal responded behaviorally.

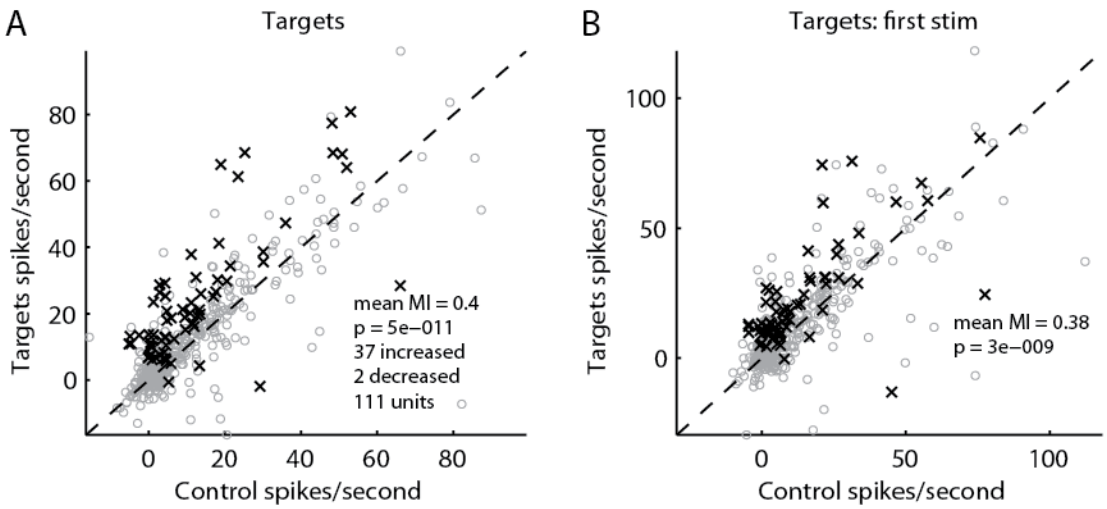


Figure 5.4. Firing rate increases were not due to stimulus order effects. (**a**) Comparisons which combined firing rates in hit and miss trials, and (**b**), comparison for first stimulus of target trials. In the first stimulus comparison (**b**), black x's represent data points which were significant in the comparisons including all target trial responses.

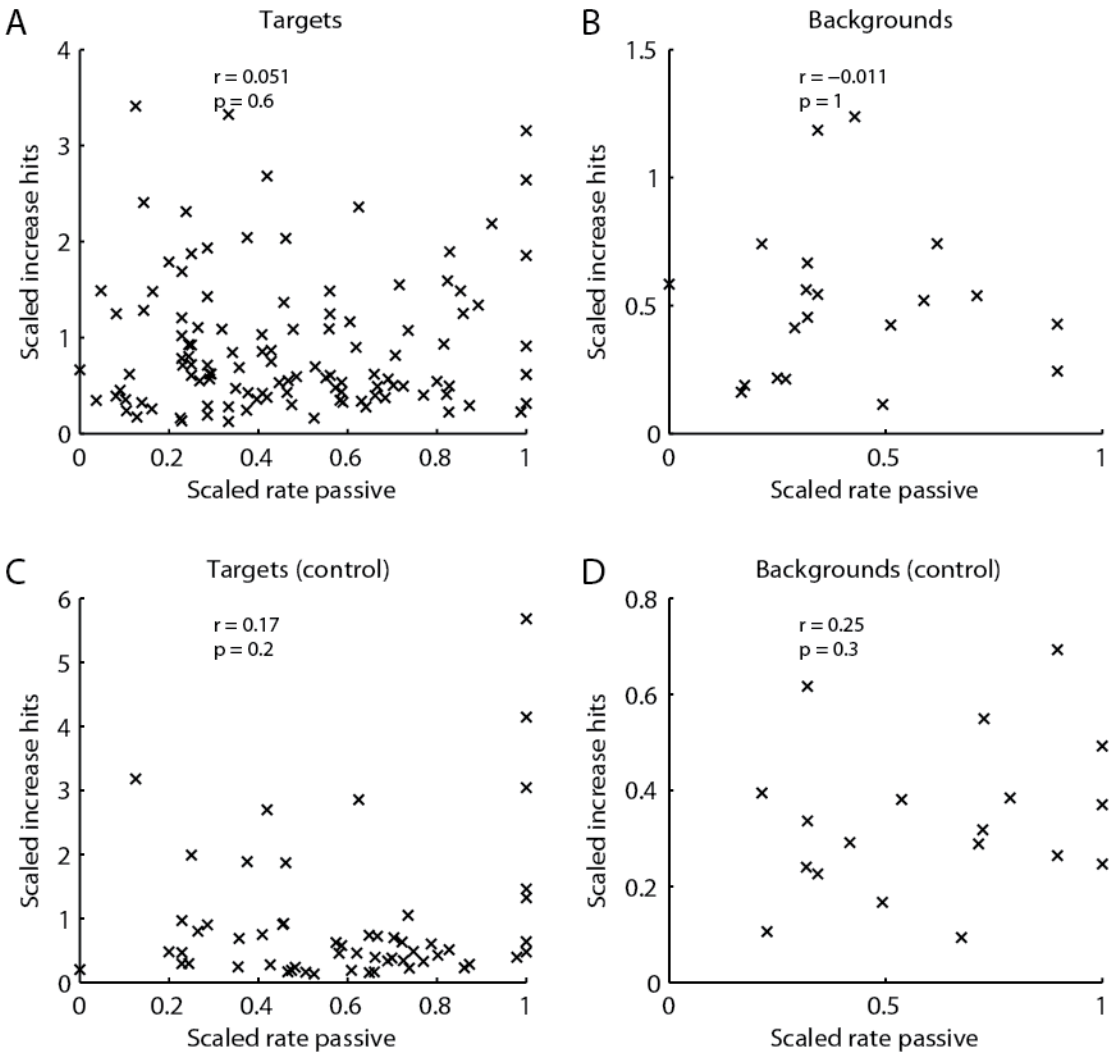


Figure 5.5. Increases occurred throughout receptive fields and were additive. (**a**) Firing rate increases plotted vs. passive firing rate for all increased data points in the behaving condition, both scaled by the maximum firing rate in the passive condition. Increases in firing rate occurred throughout space, as indicated by the large number of increases that occurred for locations which drove the neuron poorly in the passive condition. The lack of correlation in between rate increases and passive firing rates indicates that this effect was additive rather than multiplicative. (**b**) An identical analysis for increases at the background locations show the same trend, however note that throughout the population, responses to background locations did not tend to be increased. (**c,d**) Consistent results for control comparison.

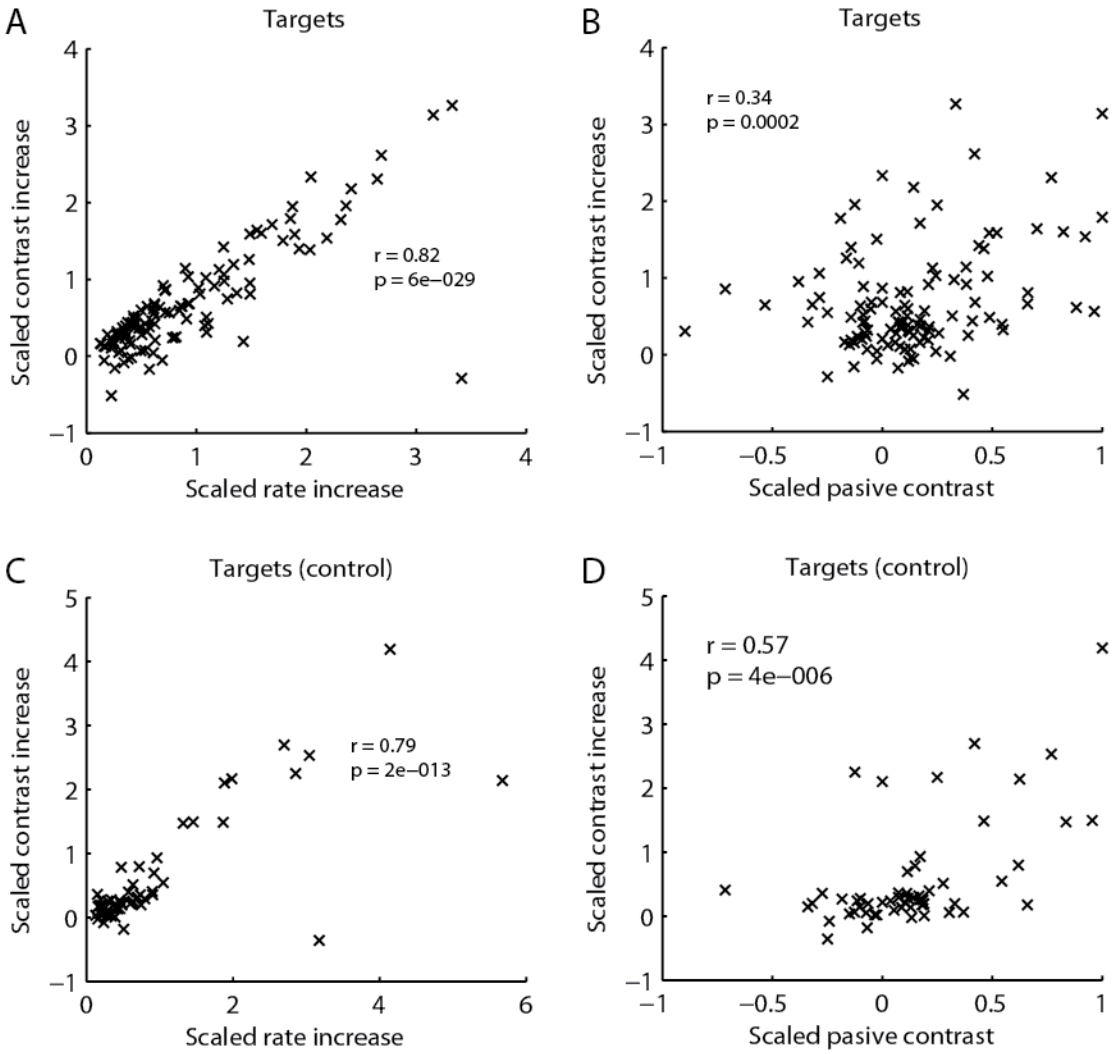


Figure 5.6. Increased firing rates to target locations increased firing rates relative to background locations. (**a**). Plot of scaled contrast increase (contrast = $R_B^T - R_B^B$) vs. scaled rate increase for target locations, showing that increased firing rates at target locations increased firing rates at those locations relative to background locations. (**b**) Scaled contrast increase vs. contrast in the passive condition. There was a weak correlation between passive contrast and contrast increase. This suggests that firing rate increases at the population level act to increase target/background contrast in neurons which already may provide a positive detection signal for target locations. (**c**) Consistent effects as in (**a**) for control comparison. (**d**) The correlation between scaled contrast increase and scaled passive contrast was slightly higher for the control comparison. This suggests that firing rate increases at the population level act to increase target/background contrast in neurons which already may provide a positive detection signal for target locations.

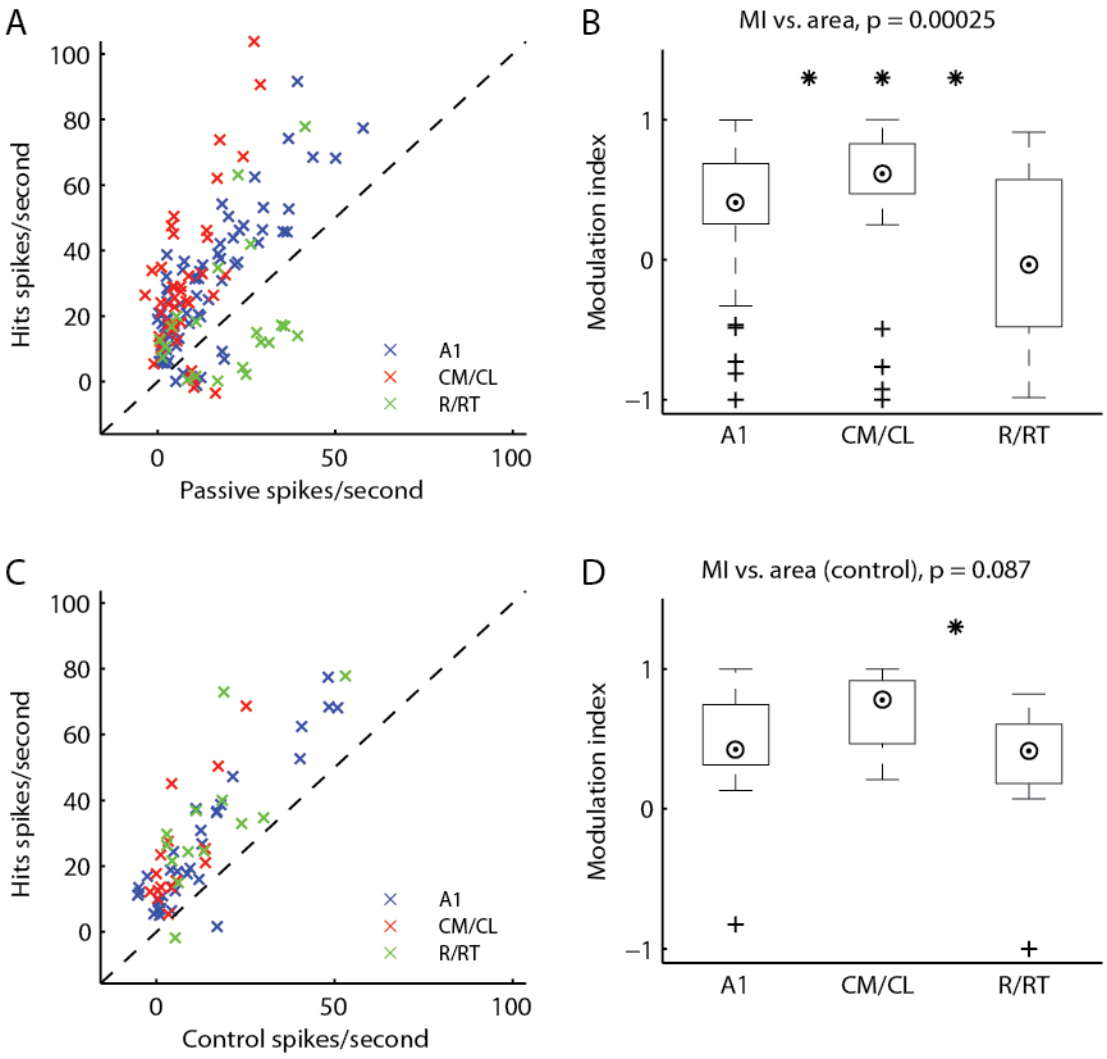


Figure 5.7. **Comparison of behavior effects in areas A1, R/RT, and CM/CL.** (**a**). Distribution of significant changes separated by area. (**b**) Effects of behavior were largest in CM/CL and weakest in R/RT. The lack of effect in areas R/RT is striking, but may be an artifact of the comparison with the passive condition, as MI was significantly greater than 0 in the behavior/control comparison. (**b,c**) The most obvious difference between **b** and **a** is the lack of significantly decreased rates in the behavior vs. control condition for all areas. Effects of behavior were largest in CM/CL and weakest in R/RT. In the control comparison, however, this effect was only significant for the comparison of CM/CL and R/RT (ANOVA, $p = 0.09$, CM/CL vs. R/RT, $p < .05$).

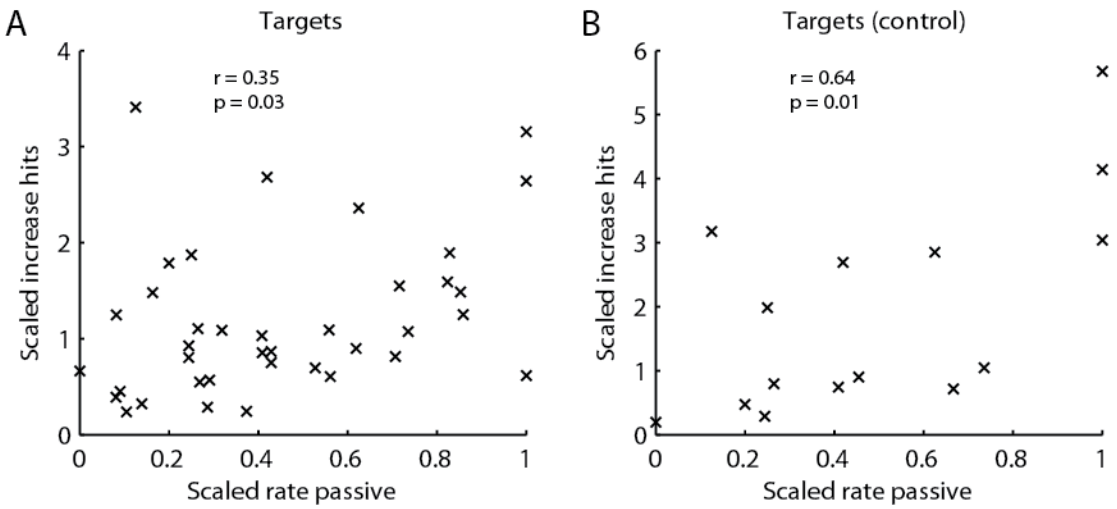


Figure 5.8. Rate increases vs. scaled rate in CM/CL. (**a**) Firing rate increase vs. passive firing rate for significant firing rate increases in CM/CL show a small correlation, indicating that effects in this area may be multiplicative and additive. (**b**) Consistent effect for control comparison.

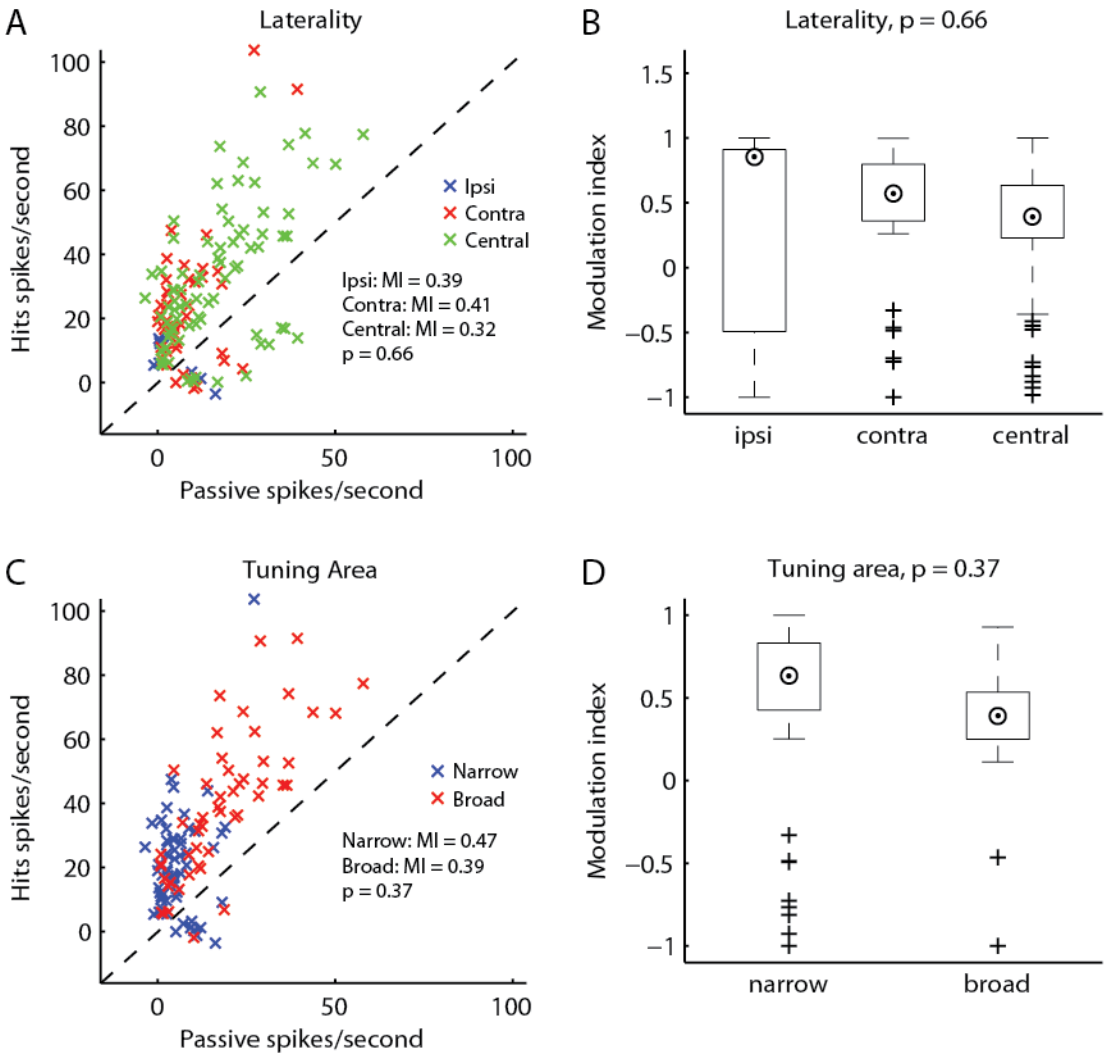


Figure 5.9. Behavior effects are not strongly dependent on spatial tuning properties. A significant effect of laterality (**a,b**) or tuning area (**c,d**) was not observed.

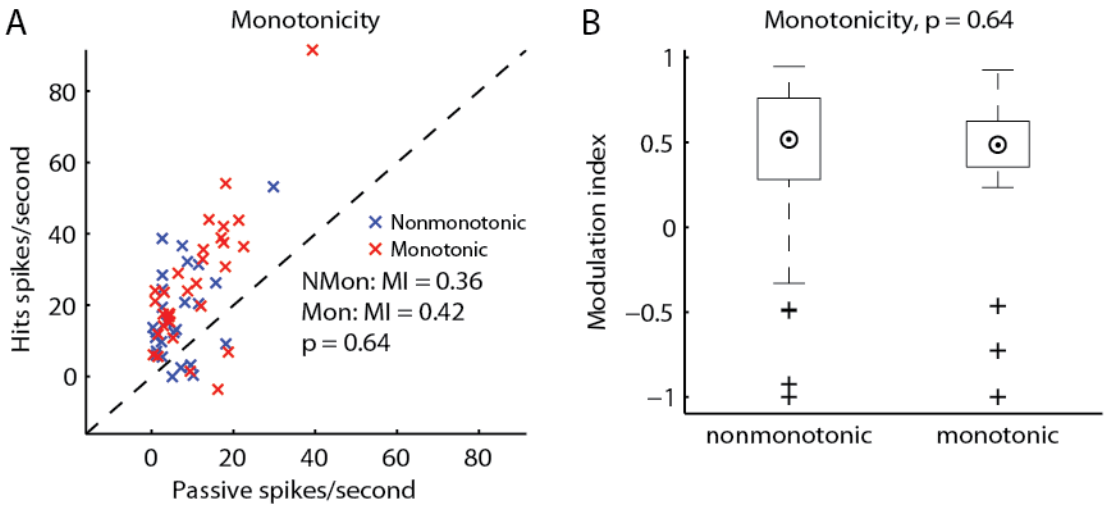


Figure 5.10. Behavior effects did not depend on level tuning. A significant effect of monotonicity was not observed.

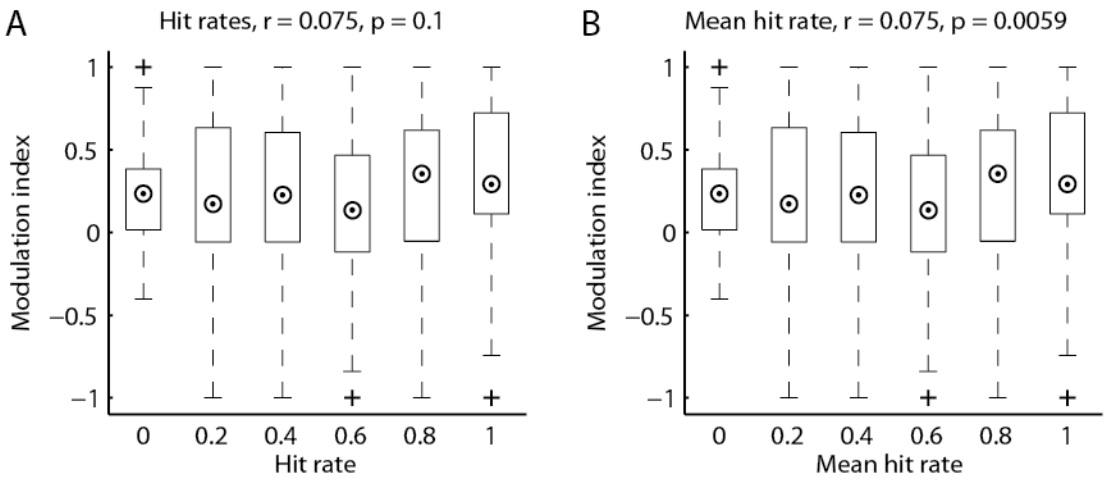


Figure 5.11. Discrimination performance was a poor predictor of behavior modulation. Behavior performance was variable between locations and behavior sessions, although behavior sessions in which animals were not engaged in the task due to lack of motivation were not included in analyses. Individual (**a**) and session average (**b**) hit rates were poor predictors of the effects of behavior on firing rates.

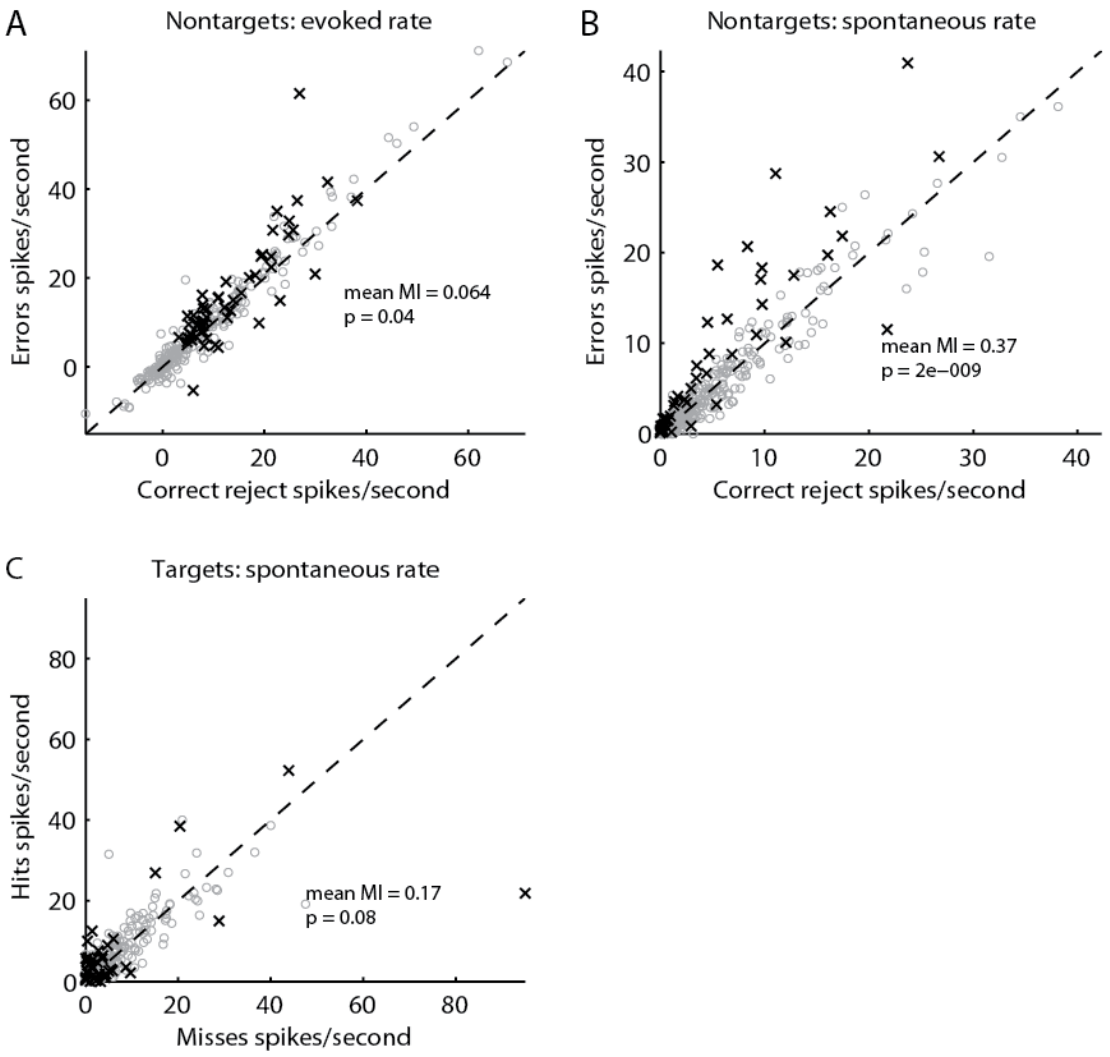


Figure 5.12. Evoked and spontaneous firing rates to background locations (**a**) A barely significant evoked increase to background locations in error trials vs. correct rejections was observed, although there was a substantial increase in spontaneous firing rate during such trials (**b**). This effect on spontaneous firing rates was not consistent when comparing hit trials to miss trials (**c**).

Chapter 6: Miscellaneous observations

6.1 Complex rate-level functions in marmoset auditory cortex

In analyzing spatial tuning properties in chapter 3, we compared variation of spatial receptive field parameters, notably tuning width, with several non-spatial response properties. Rate level tuning has been shown to be correlated with spatial tuning in several previous studies (Imig et al. 1990; Barone et al. 1996; Stecker et al. 2003), specifically in that neurons with nonmonotonic rate level functions have narrower spatial tuning and show less tendency to increase in area as sound level increases. In fact, a higher distribution neurons in the caudal areas with nonmonotonic rate-level tuning has been suggested to underlie sharper and more level tolerant spatial tuning in the posterior auditory field in cats (Stecker et al. 2003). We decided to test for this possibility in awake marmosets. However, the measure of monotonicity for cortical neurons used in previous studies of spatial tuning:

$$MI = \frac{R_{\mathrm{SPL\,max}}}{\max(R)} \tag{6.1}$$

where $R_{\mathrm{SPL\,max}}$ is the rate at the highest sound level tested and $\max(R)$ is the peak firing rate across all sound levels, seemed to insufficiently describe the behavior of a significant portion of the neurons observed. Specifically, if a neuron had a local, but not global maximum in its rate level function below the peak sound level, which is by the mathematical definition "nonmonotonic," its monotonicity index of 1given by equation 6.1 would indicate the rate level function was purely monotonic. Figure 6.1.1-2 shows several examples of such complex neurons which were clearly nonmonotonic, yet would not be classified as such using the classic monotonicity index. To better quantify monotonicity in the present dataset, a modified monotonicity index was used:

$$MI = \frac{R_{SPL\,max}}{R_{SPL\,max} + R_{local\,max}} * \frac{R_{\max\,(SPL)}}{\max\,(R)} + \frac{R_{local\,max}}{R_{SPL\,max} + R_{local\,max}} * \frac{R_{local\,min}}{R_{local\,max}} \quad (6.2)$$

This augmented modulation index is a weighted average of the classic modulation index and the ratio between the first local maximum ($R_{local\,max}$) in the rate level function and the following local minimum ($R_{local\,min}$). The weighting prevents the modulation index from being dominated by local maxima when such maxima are small and the global maximum is at the maximum sound level. The criterion for a response to be local maximum is that the neuron must be significantly driven at that sound level and the response at the next highest sound level must be smaller. For neurons with only a single global maximum (at the highest sound level or otherwise), the augmented and classic monotonicity index measures will be identical. Figure 6.1.1-2 shows several neurons and the resulting augmented modulation index. Figure 6.2A-C shows population distributions for classic monotonicity index, the local min/max ratio, and the augmented monotonicity index. Rate level functions with local min/max ratios smaller than the classic monotonicity index by 0.2 or more were considered to be "complex." Out of 487 units tested, 77 neurons reached this criterion, indicating that such multi-peaked responses represent a significant population in auditory cortex. Although there was not a significant effect of area on monotonicity in general, the smallest fraction of neurons with complex rate-level tuning was in R/RT (6/104), followed by A1 (44/300), and CM/CL which had the highest fraction (28/84). CM/CL also had the largest difference between classic and augmented monotonicity index (mean classic = 0.58 vs. augmented = 0.49).

Such complex rate level functions have been observed previously in auditory cortex (Pfingst & O'Connor 1981) as well as subcortical areas (Stabler 1996). It is unclear whether intervening studies measuring level tuning observed such complex response patterns. If they did not, their use of the classic monotonicity index would have been appropriate. It is not immediately obviously, however, why there should be such a discrepancy. One potential explanation is that complex level tuning is less common in auditory cortex of anesthetized

animals, as many studies of auditory cortex over the past several decades have been carried out in anesthetized animals. However, a recent study conducted in awake marmosets in our laboratory did not report the existence of complex rate level functions (Sadagopan & Wang 2008). Another potential explanation is the fact that the majority of studies measuring rate level functions, or at least those in awake animals, have measured level tuning to pure tones, whereas the present data are primarily responses to broadband stimuli. However, this also may not by itself be a sufficient explanation, as the existence of complex rate level functions (25% of neurons reported) was originally reported in response to pure tones (Pfingst & O'Connor 1981). It is possible that differences could be due a limitation on the highest sound level played in different studies. In the previous study in our laboratory, the maximum sound level tested was 80 dB SPL (Sadagopan & Wang 2008), whereas in the macaque study, units were tested at least up to 90 dB SPL (and in some cases above 110 dB SPL; see figure 5 in Pfingst & O'Connor (1981). In the present study, sound levels were limited to 85dB SPL (equivalent pure-tone peak to peak amplitude), though it is not straightforward to compare sound level in the auditory system for narrowband and broadband sounds due to the filtering of the cochlea. Finally, this study and the Pfingst & O'Connor (1981) study primarily used stimulus durations of 200 milliseconds, whereas Sadagopan & Wang (2008) used 100 ms sounds. If the complex behavior is related to changes in sustained firing rather than onset responses, longer duration sounds could increase the incidence of complex rate level functions. We therefore performed an analysis of the data using 100 ms, 50 ms and 20 ms post-stimulus onset windows (with a 15ms shift for response latency) to test for this possibility (figure 6.3). There was a significant increase in monotonicity as the analysis window was shortened (ANOVA, $p < 0.001$), and some units which were complex when analyzed across the full stimulus duration became either monotonic or classically nonmonotonic, particularly when using a 50 ms or 20ms analysis windows. Figure 6.4.1-3 shows several neurons analyzed at different time analysis windows to illustrate this behavior. This suggests that differing temporal response properties (i.e. onset vs. sustained responses) across sound level play a

significant role in complex rate level behavior. However, the existence of complex rate level tuning was not dependent on long analysis windows (100 ms window: 78/402 units, 50ms window: 47/273 units, 20ms window: 15/114 units). In particular, this rules out the stimulus duration explanation for the lack of complex rate level functions reported in Sadagopan et al. (2008). Therefore, it seems likely that complex rate level tuning is most common in awake animals in response to loud or broadband sounds.

6.2. Prediction of spatial receptive fields using RSS frequency weighting functions.

6.2.1. Introduction

The highly heterogeneous distribution of spatial receptive fields measured in chapter 3 warrants further investigation of how such varied receptive field shapes arise. Previous studies have shown that spatial tuning in virtual acoustic space can be predicted largely using linear binaural spectral weighting functions and head related transfer functions (Schnupp et al. 2001; Slee & Young 2013). However, spatial receptive fields in these populations were largely homogenous, confined to the frontal and contralateral hemifields, and more heavily dependent on binaural and level cues rather than spectral cues. Therefore, we attempted to predict spatial receptive field shapes using linear weighting functions generated using random spectral shape stimuli and acoustic HRTFs to see if spectral tuning properties of the neurons studied could predict some of the more interesting SRF types, such as front/rear and elevation preference.

6.2.2 Methods

Prediction of spatial receptive fields with spectral weighting functions.

Spectral weighting functions were calculated in the same manner as spatial weighting functions, using methods described in previous studies (Yu & Young 2000; Barbour & Wang 2003a). In this study, all RSS stimulus sets were constructed with a bin density of 10 bins per

octave over 2-32 kHz (4 octaves) resulting in a total of 41 spectral weights. The standard deviation parameter was 10dB and stimuli were constructed to contain 3 tones per bin. Sets contained either 43 stimuli (the minimum 42 required for an orthogonal set plus 1 flat spectrum stimulus), or 255 stimuli (3 sets and their inverses and 3 flat stimuli). For sets of 43 stimuli, 5-10 stimulus repetitions were presented; for sets of 255 stimuli, 1-2 repetitions were presented. Once weighting functions were calculated, they were used to predict responses to broad band sounds presented from different locations based on the head-related transfer function measured from each location (see below: Measurement of head-related transfer functions). The procedure for predicting spatial receptive fields was as follows: 1). Measure the actual spatial receptive field and select the location with firing rate nearest the mean firing rate across all locations to be the reference location L_0. 2). Generate a spectral level matrix Λ based on the level difference (in dB) across the sound spectrum between the HRTFs at each location and the reference location. The reference location now becomes the "flat spectrum" and the firing rate to L_0 becomes R_0 from the synthesis equation (equation 2.5). Now with Λ, R_0, and w from the RSS analysis equation, firing rates to all spatial locations can be predicted using the original synthesis equation. For each neuron, six different predictions were made, one for 5 different binaural interaction types (excitatory-excitatory (EE), contralateral-excitatory (EO), contralateral-excitatory/ipsilateral-inhibitory (EI), ipsilateral-excitatory (OE), and contralateral-inhibitory/ipsilateral-excitatory), and a final prediction using a random permutation of the frequency weighting function. As we were not able to measure binaural spectral weighting functions using free-field sound delivery, all calculations were performed as though the spectral weighting function for each ear was either identical, or the inverse of the measured spectral weighting function. The final SRF prediction was that with the highest correlation with the actual observed spatial receptive field.

Measurement of head-related transfer functions.

Head-related transfer functions were measured in a similar manner as described previously (Slee & Young 2010). However, the goal was to measure the acoustic effects of the complete experimental setup (i.e. head holder, electrode manipulator, and stereotaxic frame; Zhou & Wang 2012) in order to most accurately use these transfer functions to predict responses to spatial stimuli presented in the setup. The acoustic stimulus was a pair of 16384 point Golay code (Zhou et al. 1992) averaged over 20 repetitions. The acoustic signal was measured near the subject's tympanic membranes by placing a pair of small microphones (Knowles Electronics model FG-23329-C05, Digikey) directly into the ear canals. All subjects were head-fixed (as in recording); one animal was briefly anesthetized with ketamine/xylazine to facilitate microphone placement. The microphone signal was amplified (40dB, custom built amplifier) and digitized (RX6, Tucker-Davis technologies). The average response to the Golay codes was used to calculate the transfer function for each location (Zhou et al. 1992).

6.2.3 Results

SRFs predicted using spectral weighting functions tended to have poor correspondence with the observed spatial receptive fields. There was a subset of neurons with relatively good predictions ($r > 0.5$, n = 42/111 tested), although 4 of these neurons' had higher correlation values when using randomly permuted weighting functions, suggesting that this estimation procedure is sensitive to noise. Well predicted neurons had various spatial preferences, including the contralateral and ipsilateral hemifields, front and rear locations, and locations above and below the midline. Figure 6.5.1-5 shows several examples of neurons with good SRF predictions. The mean best prediction correlation was 0.41 (Figure 6.6A). In general, the predicted responses were highly compressed, as can be seen by plotting the actual SRF tuning areas and predicted SRF tuning areas (Figure 6.6B). The distribution of best binaural prediction type is shown in Figure

6.6C. Sufficient spectral weighting functions were not collected from neurons in areas CM/CL and R/RT to make meaningful comparisons between areas.

6.2.4 Discussion

These results show that spatial receptive fields in can be roughly predicted in a small portion of neurons in auditory cortex of an awake animal using spectral weighting functions measured in the free field. These preliminary results, however suggest that a large fraction of spatial receptive fields are not well predicted by linear weighting functions, in contrast with previous findings (Schnupp et al. 2001; Mrsic-flogel et al. 2005). Of course, this comparison must be taken with the caveat that this estimation procedure does not characterize binaural spectral weighting functions individually, but rather assumes a symmetric weighting function based on one of several binaural classes. Addressing these issues would require closed-field acoustic stimulation.

6.3 Detection of repeated motifs in randomly generated stimuli

6.3.1 Introduction

Coarsely, there are three types of long-term auditory memory: memories of what sound or sounds were heard in a specific context (i.e. “episodic” memory), memories for the association between sounds or sound attributes and the object and/or meaning associated with the sound (i.e. “semantic” memory), and the memories of sounds or sound attributes themselves (i.e. “perceptual” memory). It is this third type of memory which enables the first two types based on the incoming waveform reaching the eardrums. This, it could be argued, is the single most important and challenging task of the auditory system. Humans have a remarkable capacity to learn and recognize sound sources: it is trivial for most people to identify, for example, dozens of musical instruments, hundreds of voices, thousands of songs, and a seemingly unlimited number of naturally occurring sounds. The identification of these sound sources requires integration of

both spectral and temporal acoustic information. Currently, it is not well understood how the mammalian brain (a) recognizes learned sounds, or (b) learns to recognize new sounds.

A factor in the lack of progress in understanding the acquisition and use of auditory memory in mammals is the notorious difficulty that mammals, particularly primates, have performing long-term (or even short-term) auditory memory tasks (Stepien et al. 1960; Colombo & D'Amato 1986; Fritz, Mishkin, et al. 2005). It has even been suggested that, among mammals, the ability to commit arbitrary sounds to long-term auditory memory may be unique to humans (D'Amato 1988; Fritz, Mishkin, et al. 2005), although this is unlikely given the demonstrated ability of some mammalian species to recognize specific human words (Kaminski et al. 2004),and primates have been shown to recognize both call type (Seyfarth et al. 1980) and caller (Rendall et al. 1996), although call type recognition could be an innate ability rather than a demonstration of auditory memory. A task which could be used to study learning and expression of long-term memory in an established primate model would be a boon to the study of the formation and use of auditory memories.

Additionally, perceptual auditory memory can be confounded by semantic or expert strategies, and even humans have a difficult time quickly learning and recognizing natural sound sources and sequences (Cohen et al. 2009). However, a recent study recently demonstrated that humans quickly and subconsciously ("implicitly") formed long-term auditory memories for random sounds (noise tokens), while performing a short term memory task detecting noise tokens with repeated motifs (Agus et al. 2010). In this task, there were two types of target stimuli: regular repeated noise (RN) stimuli, which were trial unique, and "reference" noise tokens (Ref RN) targets, which were repeated throughout a testing session. Participants were not informed of the existence of the Ref RN stimuli. However, performance in the task increased rapidly for the Ref RN stimuli within testing sessions, and these performance gains persisted across multiple

week delays, indicating that perceptual learning took place. Subjects were also able identify previously learned Ref RN tokens.

The ability to detect repeated random sequences is not a new discovery and has been used to probe short-term (Guttman & Julesz 1963; Warren et al. 2001; Kaernbach 2004), as well as long-term memory (Hanna 1984; Goossens et al. 2008), at least within a testing session. The ability to discriminate repeated noise tokens from continuous noise has also been tested in several animal models, including cats (Frey 2003), gerbils (Kaernbach & Schulze 2002), and pigeons (Kretzschmar et al. 2008). The unique finding of Agus et al. (2010) was that such learning occurred without instruction (or perhaps even subject's knowledge), and persisted for long periods of time. Thus, we thought that this task was an ideal candidate for the study of long-term auditory perceptual memory in the marmoset monkey.

6.3.2 Methods

We trained marmosets to discriminate target sounds composed of repeated randomly generated sounds from continuous random sounds. The task structure was identical to that used for auditory discrimination (see chapters 2, 4, and 5), and was a go/no-go procedure. Backgrounds were continuous random sounds, and targets were repeated motifs. Two types of randomly generated sounds were used: Gaussian noise filtered between 2-32 kHz, and random frequency modulated pure tones. For Gaussian noise, repeated motifs (hereafter repeated noise, or RN) were constructed by generating a single unfiltered noise sequence with a length equal the desired period times the sampling frequency, concatenating several copies of that sequence, and applying the filter. This ensures that there are no acoustic artifacts of the concatenation procedure. The background sound (hereafter noise, or N) was simply a single Gaussian noise of the desired length. For generation of RN stimuli using random frequency modulated tones (hereafter RFM), a random number sequence was generated with length equal to the desired period times twice the

desired modulation frequency cutoff. In this study, the FM cutoff was 20 Hz. Therefore, an RFM RN stimulus with period 500ms and duration 2000ms would have an FM contour sequence of length 20 concatenated for times. This sequence was then resampled to the system sampling frequency to generate the stimulus frequency contour $f(t)$, and this contour was integrated to produce the final sound stimulus:

$$y(t) = \int_{\tau=0}^{t} f(\tau)d\tau \tag{6.3}$$

Generation of background (N) RFM sounds was similar except that a single full-length contour sequence was used rather than a concatenation of shorter sequences for the RN. All background (N) stimuli, and half of target (RN) stimuli were generated uniquely for each stimulus presentation. The "reference" target stimuli (hereafter Ref RN) were constant throughout single, and often multiple, testing sessions.

6.3.3. Results

We first investigated the ability of marmosets to detect repeated motifs of Gaussian noise. This proved possible, but extremely difficult for the marmosets. Of three tested, only one (M94W) was trained to perform the task reliably, even for very short motif periods. The failure to train the other two was potentially due to the fact that the animals had been extensively trained in the sound location discrimination task and had difficulty generalizing. However, it was also clear that the task itself was difficult for marmosets. To study auditory memory rather than periodicity detection, it is preferable to use RN sounds with periods of at least 250 ms (Guttman & Julesz 1963; Kaernbach 1993; Warren et al. 2001). M94W, however, showed extreme difficulty in performing the task with an RN period of 200, and performance degraded further beyond that point (figure 6.7). In an attempt to reduce task difficulty, the Gaussian noise sequence with was replaced by a pure tone with a random frequency modulation (RFM). This proved to be significantly easier for the subject to discriminate; within just a few sessions, performance had

improved substantially at RN periods for which detection rates for Gaussian noise RNs had been negligible. M94W quickly learned to detect sequence repetition periods of up to and over 500ms.

Two strategies were used to show evidence for perceptual learning Ref RN tokens. First, by testing the subject with the same Ref RN stimulus in multiple sessions, a listener may increase performance over time to the Ref RN relative to trial unique RN stimuli. Task difficulty was also increased over the sessions by increasing RN period. As observed in humans detecting repeated noise motifs (Agus et al. 2010), some Ref RN tokens were better discriminated than others. However, for this subject, behavior performance did not appear to increase for Ref RN sounds which were not initially identifiable. Figure 6.8 shows performance across multiple sessions for several Ref RN tokens. Due to the difficulty of the task, we did not extensively test Ref RN tokens for which performance remained poor for more than 2 or 3 sessions.

Second, we measured the probability of a correct Ref RN detection vs. trial number across multiple sessions. In the human Ref RN study, increased performance indicative of Ref RN learning occurred within the first few trials (Agus et al. 2010). However, this analysis also indicated that Ref RN detection performance was a function of the specific Ref RN motif rather than any perceptual learning. In the 5/9 Ref RN tokens tested for which the subject displayed good performance (a higher hit rate than for the unique RN targets), the average hit rate for the first target presentation was 1, and for the 4 Ref RN tokens for which performance was poor, the hit rate for the first target presentation was 0.5.

6.3.4. Discussion

Although we did not find evidence for long-term auditory memory in the present results, we still feel that the task is promising for the study of auditory memory. First, the subject was able to detect repetitions of random sound sequences with periods of up to 1000ms, which is substantially higher than has been shown previously in any non-human mammal (Kaernbach &

Schulze 2002; Frey 2003). This is likely due to the use of the RFM stimuli, but could be indicative of the auditory abilities of marmosets. Even without a long-term memory component, this makes the task useful for the study of short-term auditory memory. Second, it is possible that the experimental design employed simply was not optimal for detecting learning over short time scales observed previously in humans. In a Go/No-Go task, subjects are rewarded only for responding to target stimuli, and not for withholding responses to non-target stimuli. An "overly eager" subject might have a shifted response bias (Swets 1961) at the start of a session, making it difficult to quantify perceptual performance during this time. A methodological improvement might therefore be the use of a two-alternative forced choice (2AFC) procedure in which subjects are required to make a behavioral response regardless of the stimulus type and can be rewarded on each stimulus delivery.

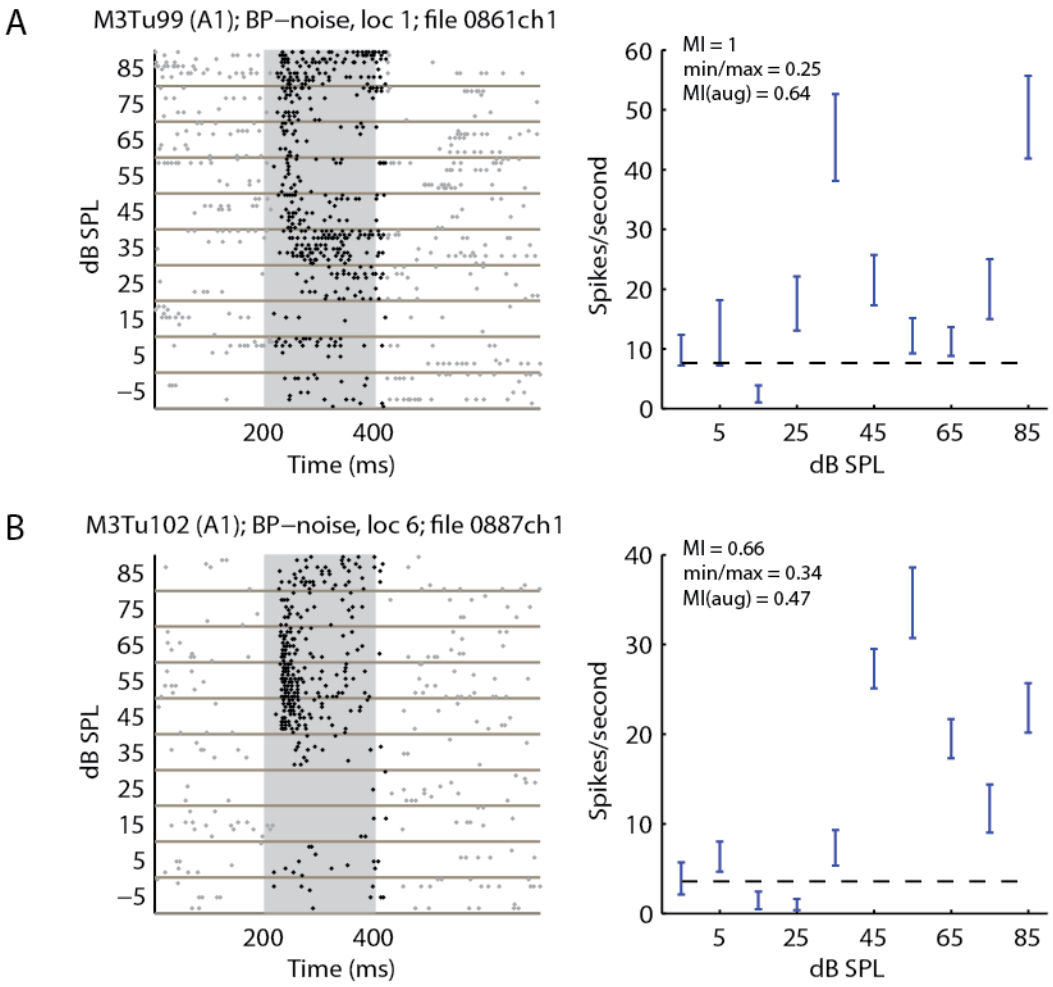


Figure 6.1.1 Examples of complex level tuning (A1). A significant portion of units exhibited nonmonotonic behavior which was not well quantified by the ratio of the firing rate at the highest level and the maximum firing rate (MI). "min/max" is the ratio of the local minimum and local maximum; "MI(aug)" is the average of MI and min/max weighted by the relative firing rates at the local maximum and the highest level. (**a**) A unit which would have been classified as "fully monotonic" (MI = 1) (**b**) A unit which would have been classified as somewhat monotonic.

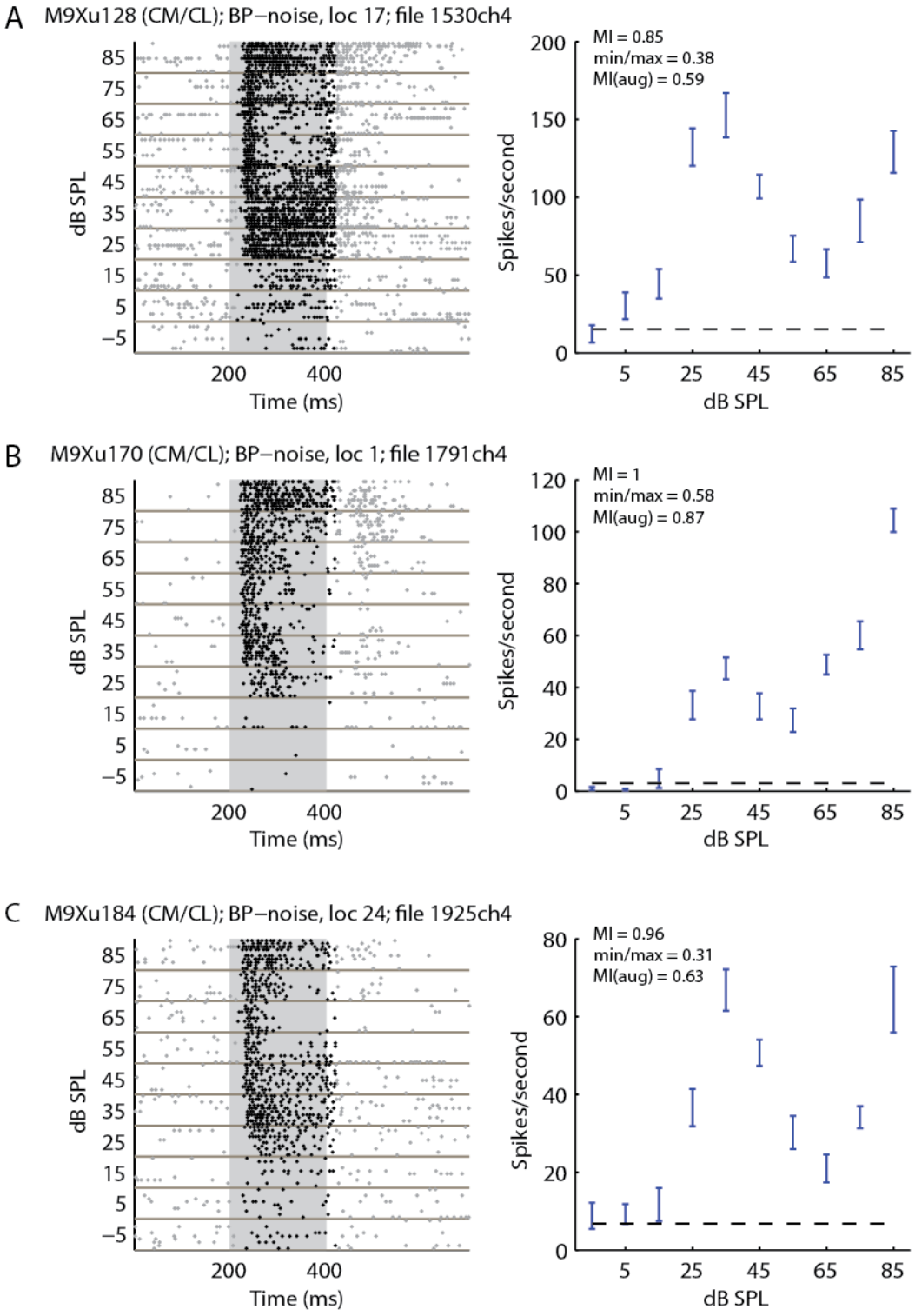


Figure 6.1.2 Examples of complex level tuning (CM/CL). (**a-c**) Units in CM/CL displaying complex level tuning.

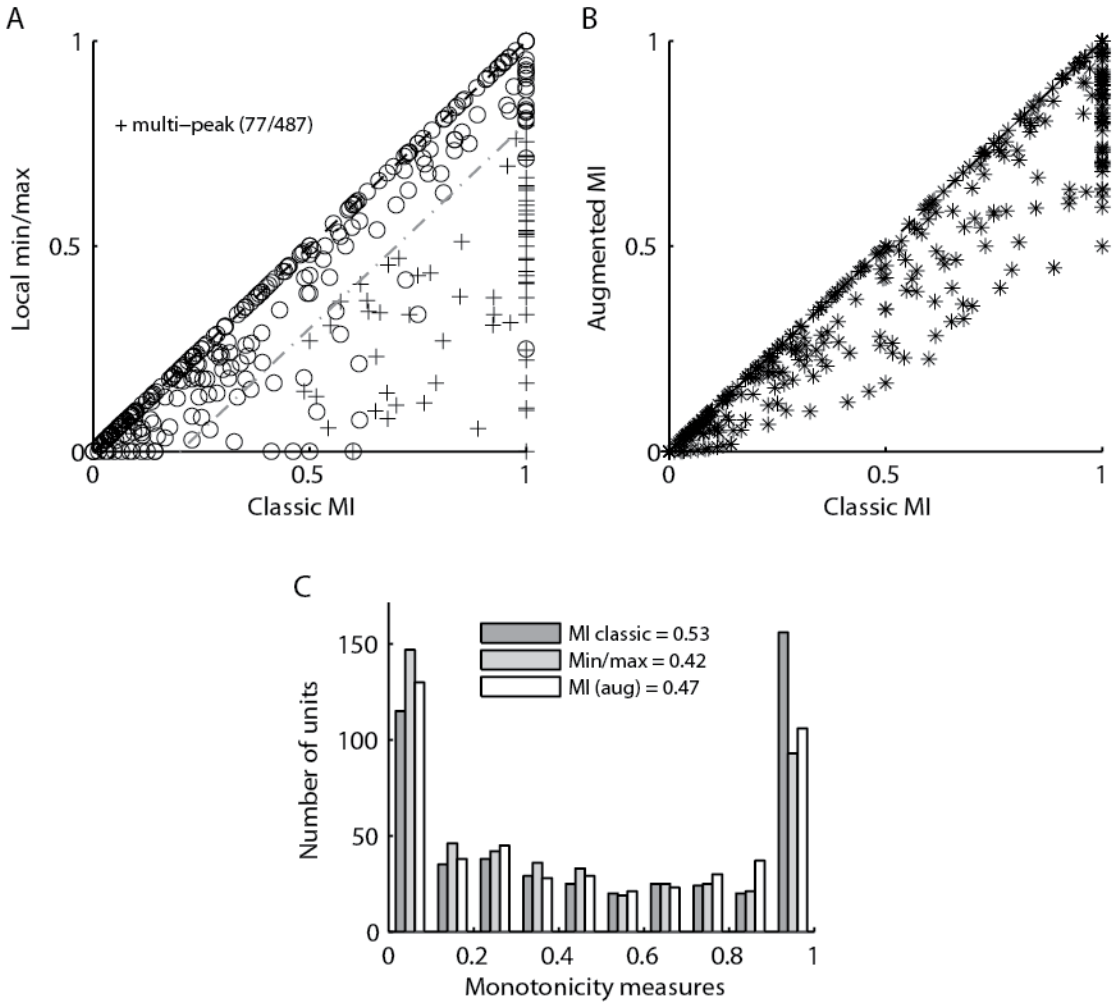
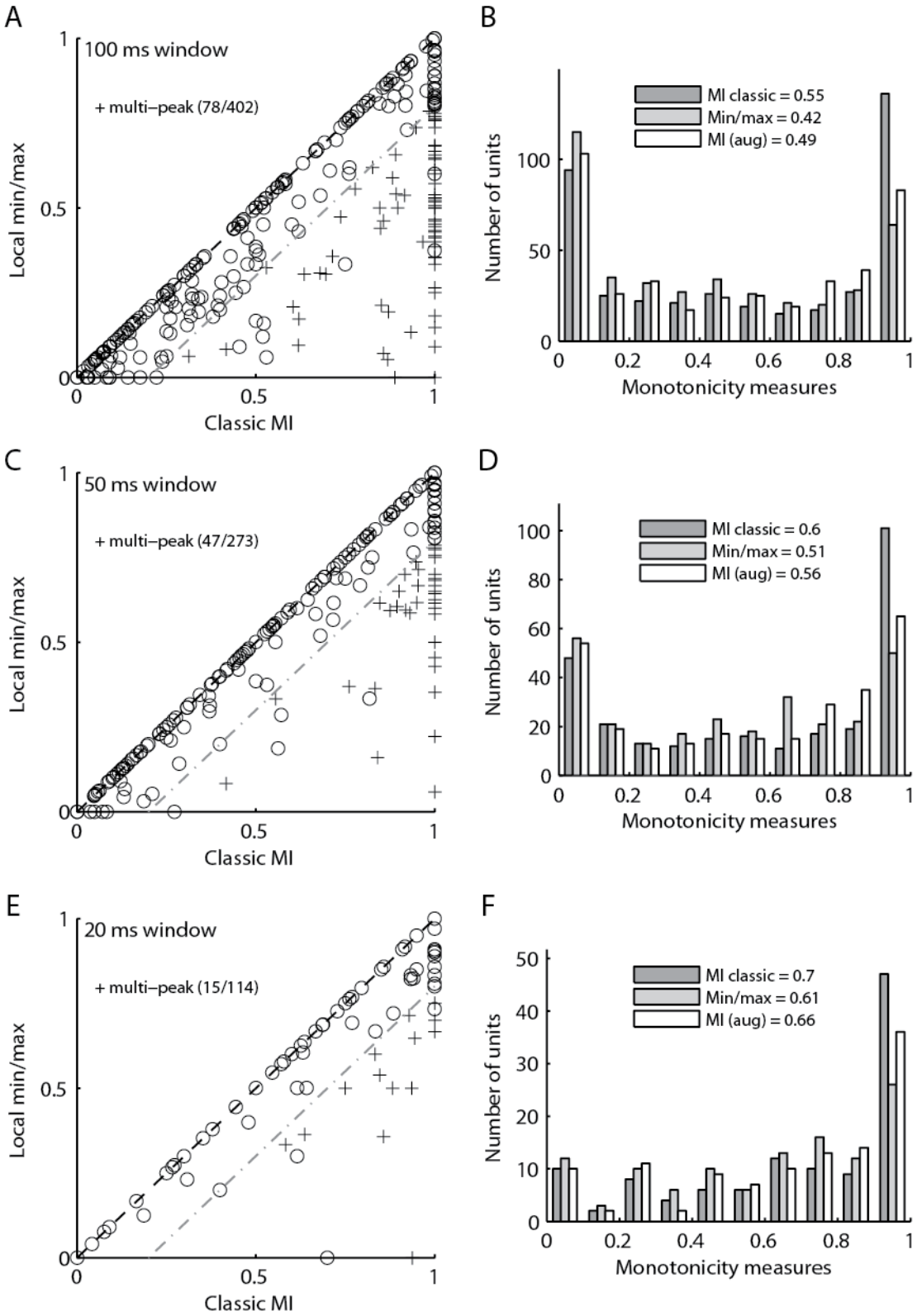


Figure 6.2. Comparison of distributions of monotonicity measures. (**a**) In a subset of neurons, the local min/max ratio was markedly lower than the classic monotonicity index, indicating that neurons with complex level tuning functions made up a significant portion of the sample neural population. (**b**). Augmented MI vs. Classic MI for the entire sample. (**c**) Population histogram for the three measures. The use of the augmented modulation index does not severely increase the degree of nonmonotonicity in the population.

Figure 6.3. Complex rate level tuning was not dependent on sustained responses. (a-d) Although monotonicity was slightly stronger at shorter analysis durations (ANOVA < 0.01), we found no apparent difference in the prevalence of complex rate level functions.


A
100 ms window
+ multi–peak (78/402)
Local min/max
Classic MI
B
MI classic = 0.55
Min/max = 0.42
MI (aug) = 0.49
Number of units
Monotonicity measures
C
50 ms window
+ multi–peak (47/273)
Local min/max
Classic MI
D
MI classic = 0.6
Min/max = 0.51
MI (aug) = 0.56
Number of units
Monotonicity measures
E
20 ms window
+ multi–peak (15/114)
Local min/max
Classic MI
F
MI classic = 0.7
Min/max = 0.61
MI (aug) = 0.66
Number of units
Monotonicity measures

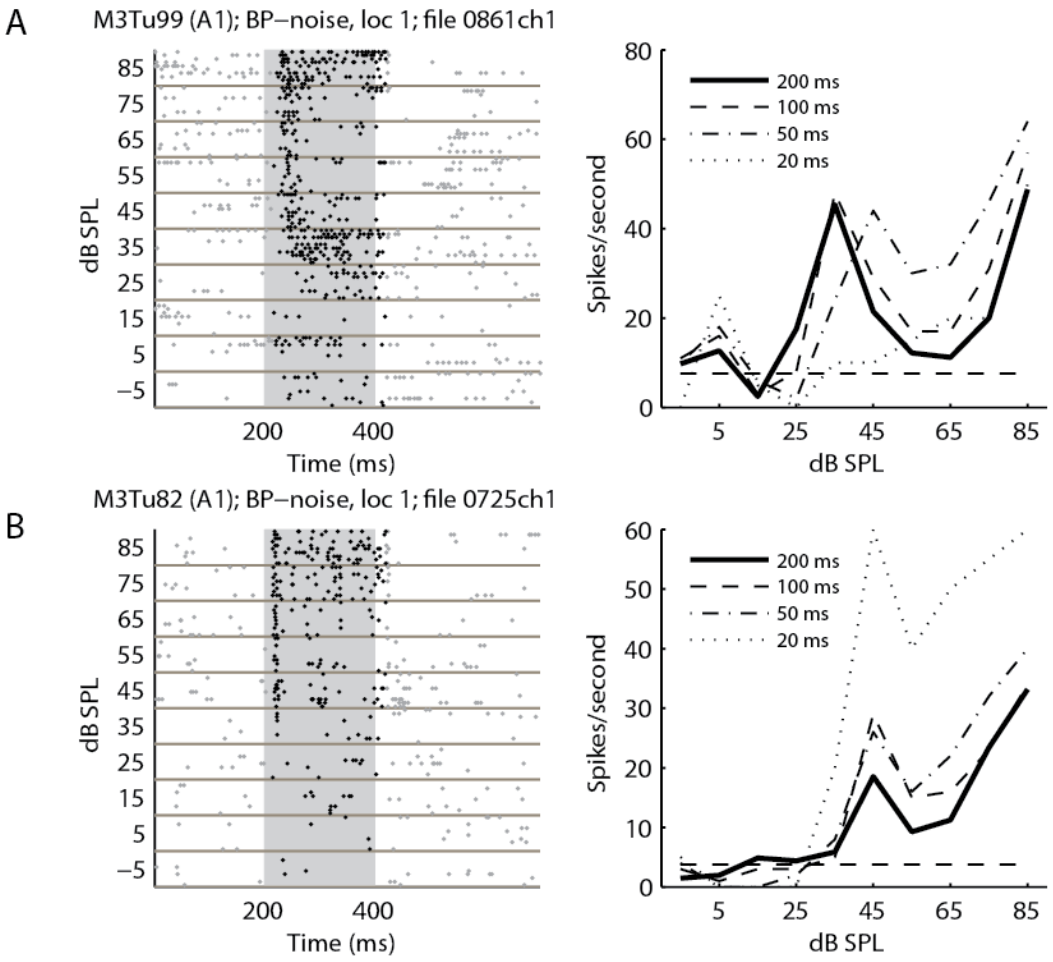


Figure 6.4.1. Rate level tuning analyzed over varying duration analysis windows. (**a**,**b**) examples of units which maintain complex rate level tuning when the time integration window for calculating firing rate is decreased from 200 to 20ms. All analysis windows begin 15ms post stimulus onset. In (**a**) the neuron is not strongly driven in the 20ms window but may be considered to be monotonic. The neuron in (**b**), however, has a strong onset response.

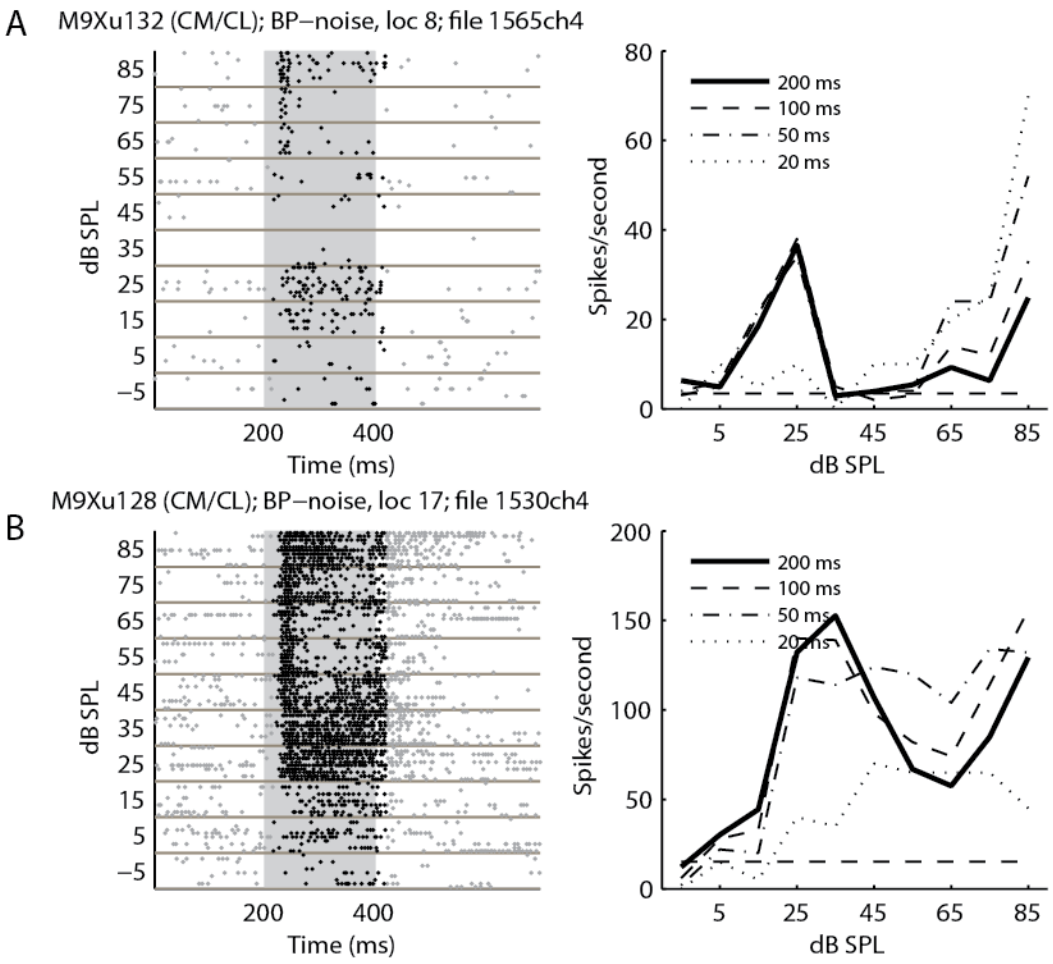


Figure 6.4.2. Rate level tuning analyzed over varying duration analysis windows (2). (**a,b**) examples of units whose rate level tuning profiles become more monotonic as the analysis window duration decreases. In these neurons, the complex rate level functions seem to be determined by the different temporal response properties at different time periods.

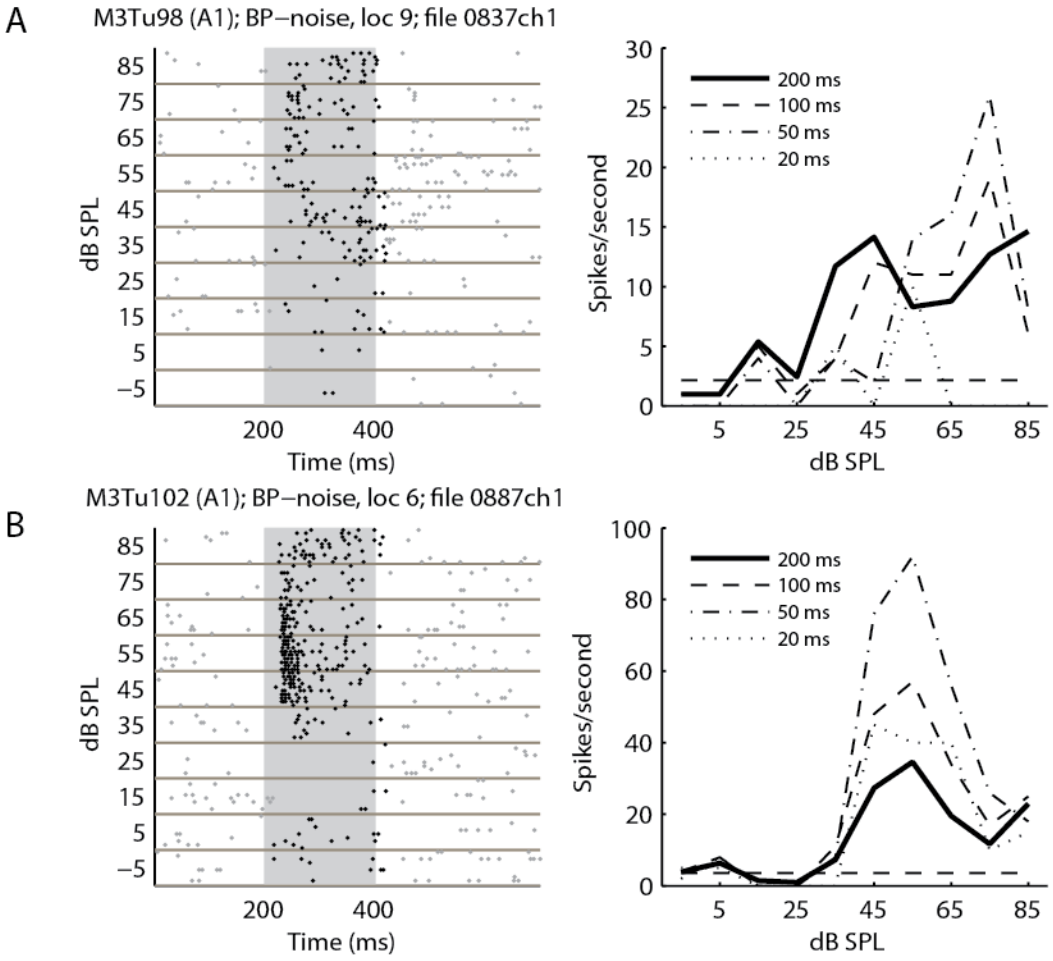


Figure 6.4.3. Rate level tuning analyzed over varying duration analysis windows (3). (**a,b**) examples of units whose rate level tuning profiles become more classically nonmonotonic as the analysis window duration decreases. In these neurons, the complex rate level functions seem to be determined by the different temporal response properties at different time periods.

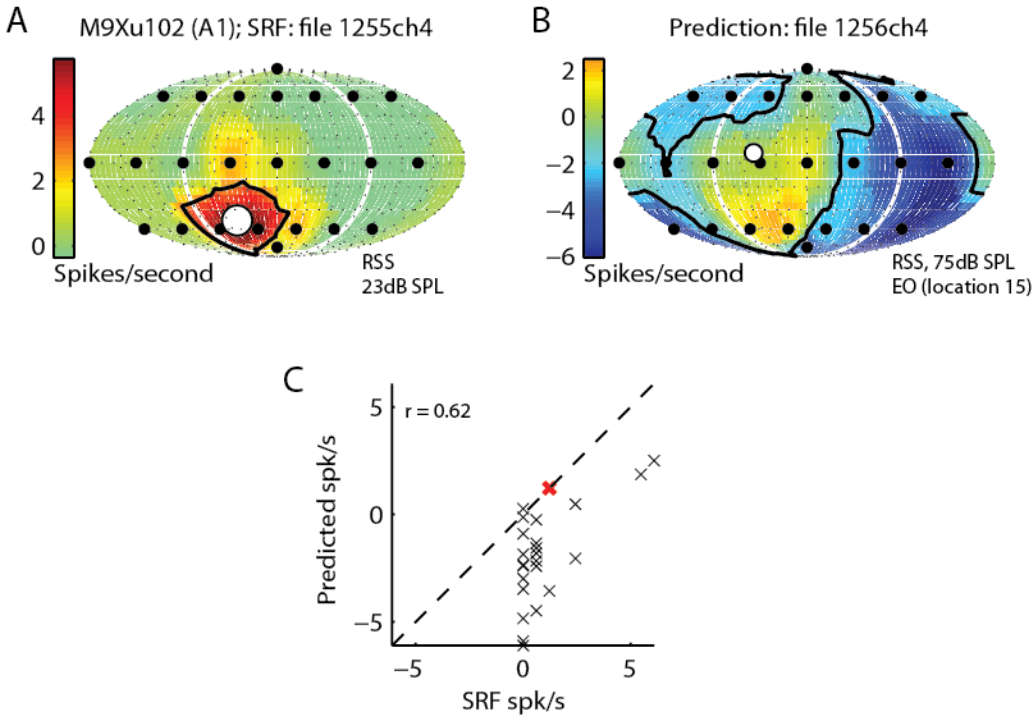


Figure 6.5.1. Prediction of SRFs using spectral weighting function (1). This neuron was among the small population of neurons recorded which showed a relatively high level of correlation between the actual spatial receptive field (**a**) and predicted spatial receptive field using the spectral weighting function and head-related-transfer function (HRTF). This neuron shows tuning below the horizon in both the actual and predicted receptive fields. The red "x" in (**c**) represents the mean firing rate location used as the reference location L_0.

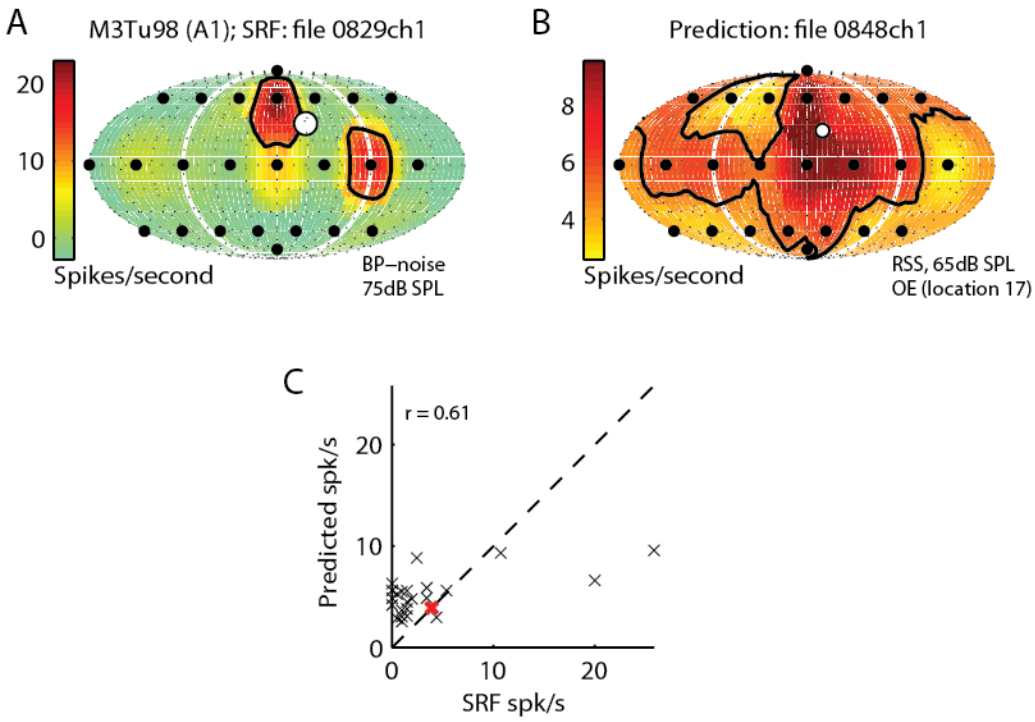


Figure 6.4.2. Prediction of SRF using spectral weighting function (2). A second example neuron which showed a relatively high level of correlation between the actual spatial receptive field (**a**) and predicted spatial receptive field using the spectral weighting function and head-related-transfer function (HRTF). This neuron shows tuning above the horizon in both the actual and predicted receptive fields. The red "x" in (**c**) represents the mean firing rate location used as the reference location L_0.

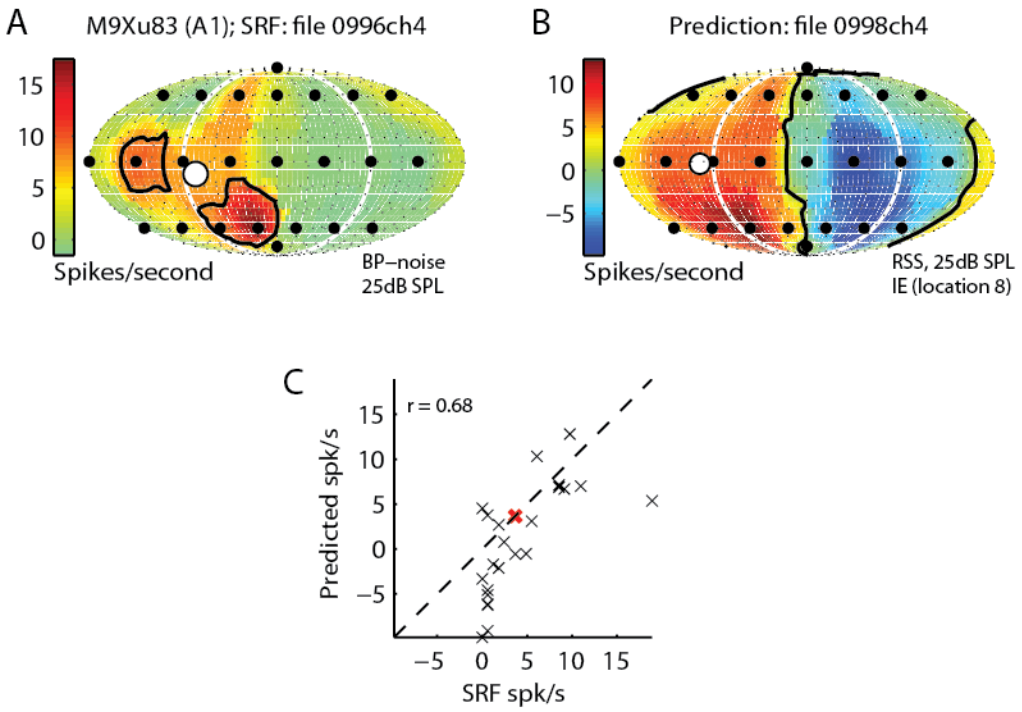


Figure 6.4.3. Prediction of SRF using spectral weighting function (3). A third example neuron which showed a relatively high level of correlation between the actual spatial receptive field (**a**) and predicted spatial receptive field using the spectral weighting function and head-related-transfer function (HRTF). The red "x" in (**c**) represents the mean firing rate location used as the reference location L_0.

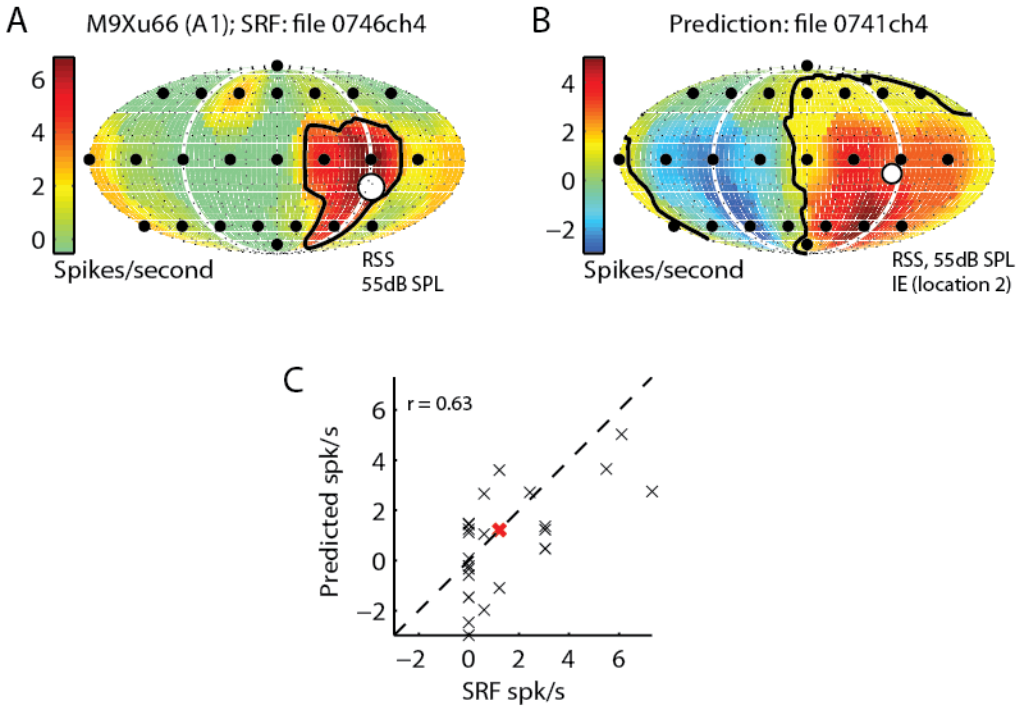


Figure 6.4.4. Prediction of SRF using spectral weighting function (4). A fourth example neuron which showed a relatively high level of correlation between the actual spatial receptive field (**a**) and predicted spatial receptive field using the spectral weighting function and head-related-transfer function (HRTF). The red "x" in (**c**) represents the mean firing rate location used as the reference location L_0.

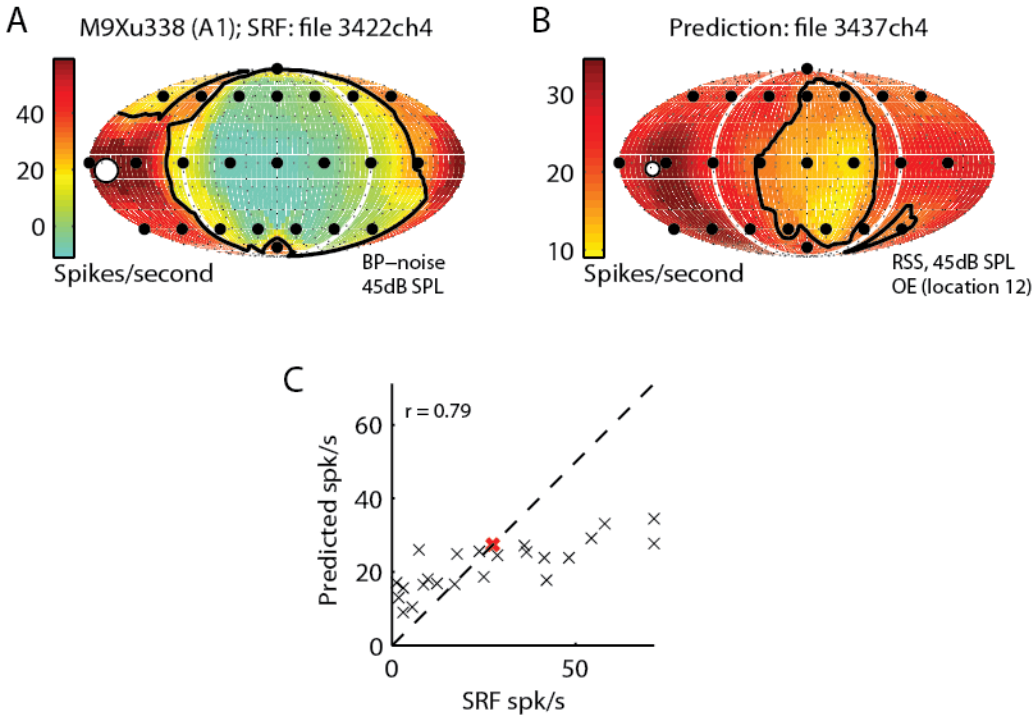


Figure 6.4.5. Prediction of SRF using spectral weighting function (5). A final example neuron which showed a relatively high level of correlation between the actual spatial receptive field (**a**) and predicted spatial receptive field using the spectral weighting function and head-related-transfer function (HRTF). This neuron is displays preference for rear locations in both the actual and predicted receptive fields. The red "x" in (**c**) represents the mean firing rate location used as the reference location L_0.

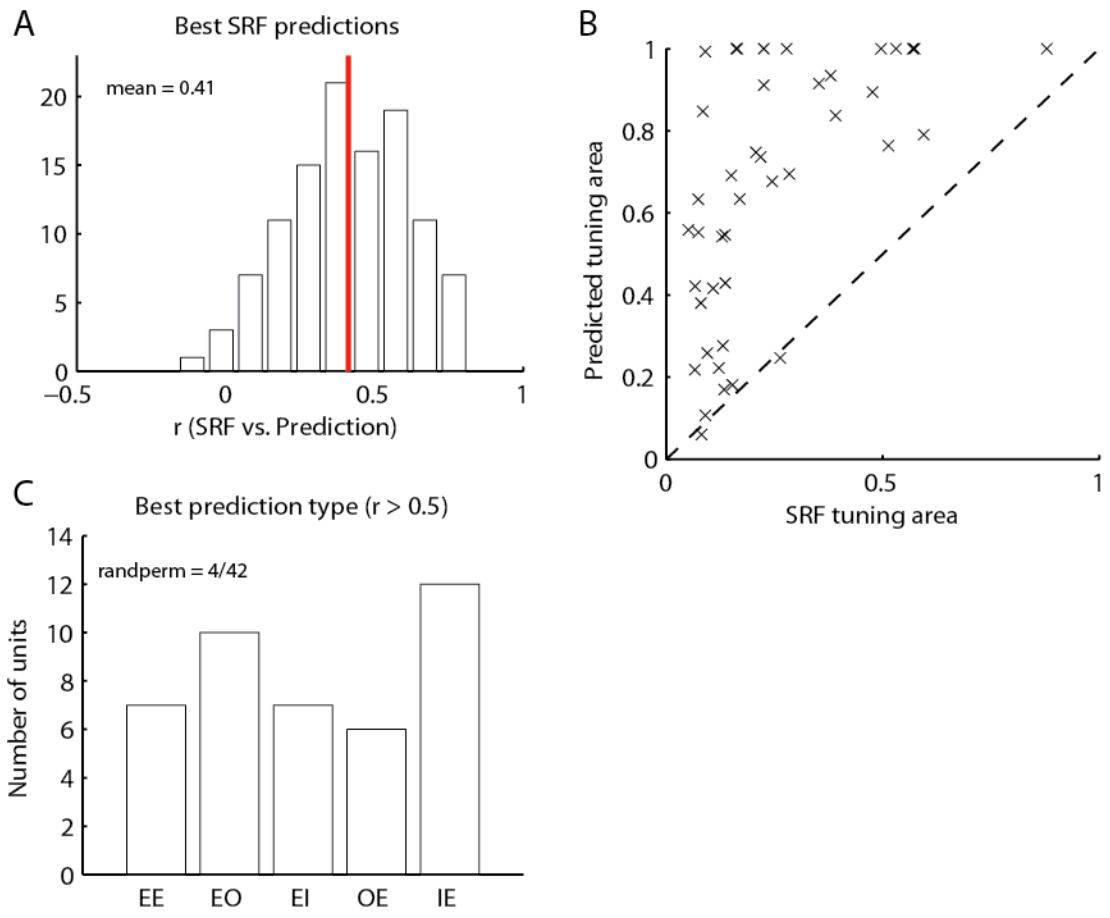


Figure 6.6. Population results of SRF predictions. (**a**) Best prediction values for all neurons tested. (**b**). Prediction tuning area vs. actual SRF tuning area in well predicted neurons (r > 0.5). Expanded tuning areas in predicted receptive fields was due to highly compressed predicted response functions. (**c**) Distributions of best binaural type for well predicted neurons. A minority of well predicted neurons were actually better predicted by randomly permuting the spectral weighting function prior to receptive field prediction.

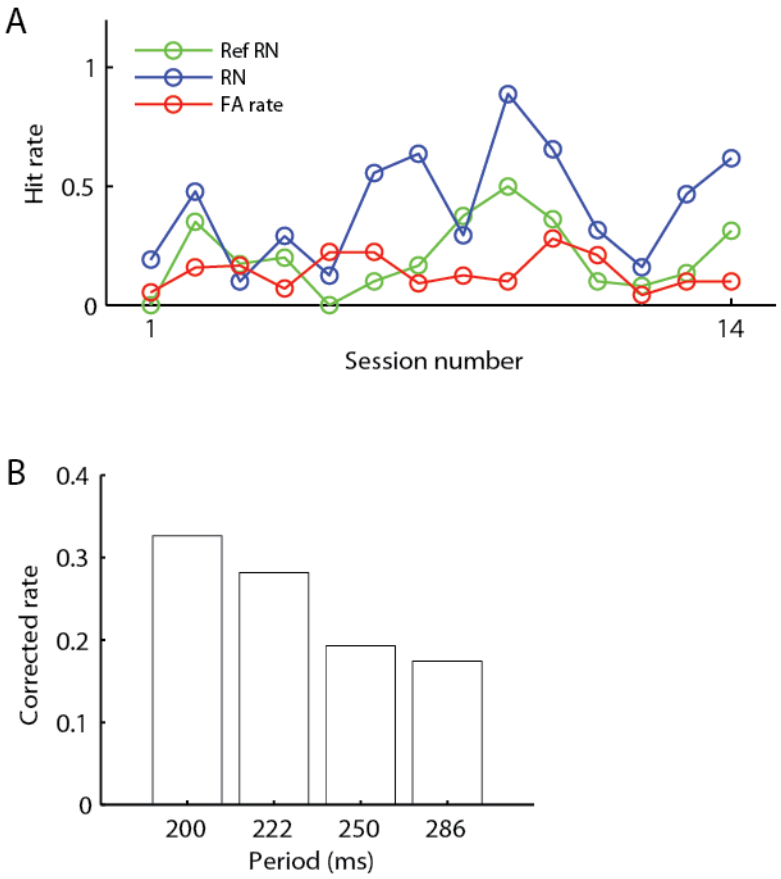


Figure 6.7. Poor detection of repeated motifs in Gaussian noise. Performance over several trials was poor and highly variable with Guassian noise stimuli when the repeated noise periods of hundreds of milliseconds. (**a**) Performance over 14 sessions with an RN period of 250ms. (**b**) Average performance over all sessions for RN periods of 200ms or greater.

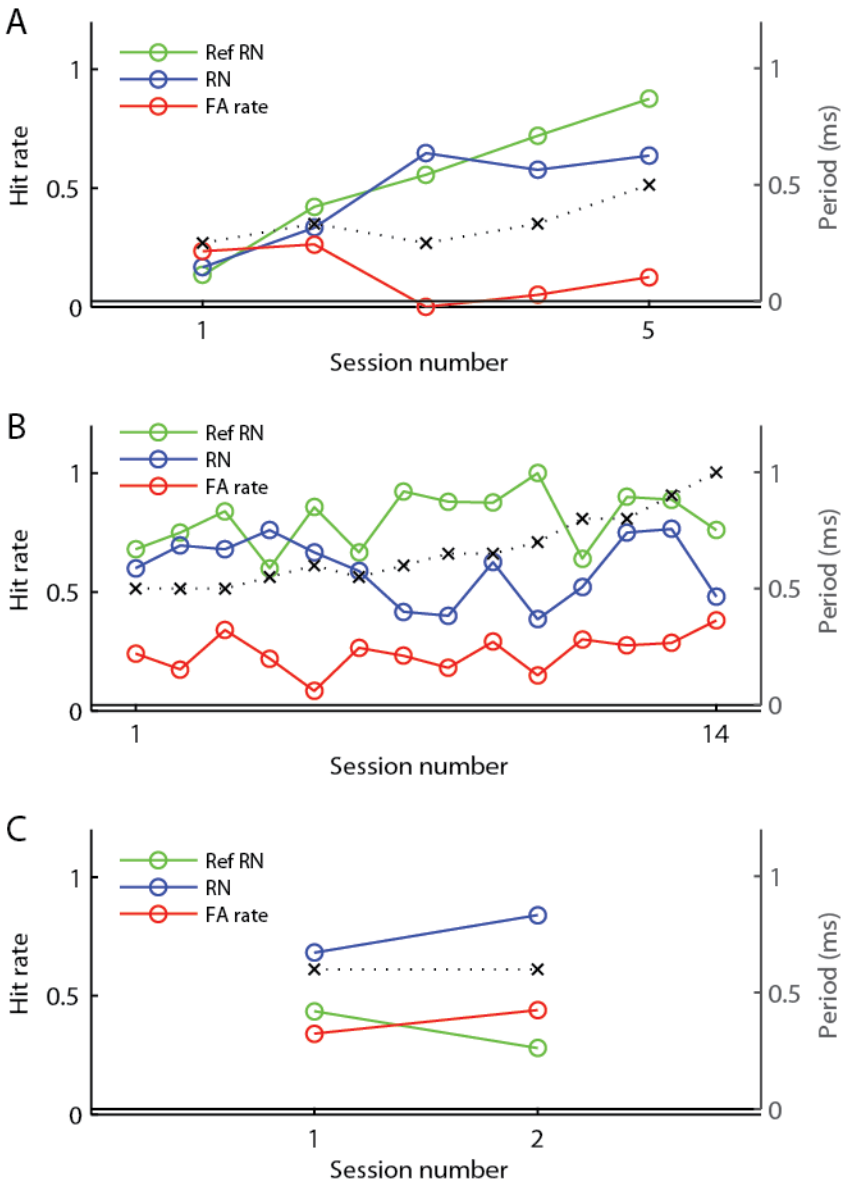


Figure 6.8. Detection of repeated motifs using randomly frequency modulated tones. (**a**) The first five sessions using RFM stimuli. The subject showed a rapid increase in the ability to detect repeated motifs in a randomly frequency modulated tone over those in Gaussian noise. Black x's denote the RN period used in each session. This subject was able to detect RFM repetitions with periods of 500-1000 ms (**b**), but did not display evidence of long-term learning of RFM motifs which were not immediately identifiable (**c**).

Chapter 7: Conclusions and future studies

7.1 Spatial receptive fields in auditory cortex

Auditory cortex, which may not directly perform computations to determine sound location, is essential for many behaviors involving sound localization in mammals (Jenkins & Masterton 1982; Bizley et al. 2007; Lomber & Malhotra 2008). How the spatial field is represented at the level of single neurons in auditory cortex however, remains incompletely understood. In this thesis, we showed that full-field spatial receptive fields in auditory cortex of an awake primate are diverse, with neurons preferring virtually all regions of space, and that these receptive fields, even those which displayed non-classical response types, are stable across sound level. This is in sharp contrast to full-field studies in anesthetized animals which consistently observed more homogeneous populations of broadly tuned neurons and whose selectivity decreased with increasing sound level in a linear and predictable manner (Brugge et al. 1994; Brugge et al. 1996; Schnupp et al. 2001; Mrsic-flogel et al. 2005). However, given some knowledge about the differences in responses between awake and anesthetized animals, perhaps it should not be so surprising. First, it has been shown in many studies of anesthetized animals that frequency tuning in auditory cortex is not invariant to sound level, becoming quite broad at high sound intensities (Heil et al. 1992; Kilgard & Merzenich 1999; Schreiner et al. 2000). However, a recent study in awake marmosets in our laboratory observed that a large proportion of neurons in primary auditory cortex were narrowly tuned to sound frequency and that this tuning was invariant to sound level (Sadagopan & Wang 2008). A related point is that many studies in anesthetized animals have found that in the majority of neurons, firing rates increase monotonically with sound level. The prevalence of nonmonotonic tuning in anesthetized studies, however, is highly variable, and in general direct comparisons between studies can be difficult to make. This is

potentially due to the different types of anesthesia used, and may also be a result of inconsistent cutoffs for monotonicity classification (Phillips & Orman 1984; Barone et al. 1996; Schreiner et al. 2000; Stecker et al. 2003). In awake animals, the prevalence of nonmonotonicity has been shown to be as high as 81% (Pfingst & O'Connor 1981), although this study had a relatively high threshold for monotonicity (0.9). There are two notable studies which explicitly reported the prevalence of nonmonotonicity and also measured spatial tuning. In the first, the mean nonmonotonic strength was 35.3%, equivalent to a monotonicity index of 0.65 (Barone et al. 1996), higher than the present mean monotonicity index of 0.48 in A1. In the second, only 9% (10/117) of units in A1 had a monotonicity index of less than 0.5 (Stecker et al. 2003), considerably lower than the fraction of 54% (163/300) observed here. Given that these studies also found that spatial tuning in nonmonotonic units was narrower and less sensitive to sound level than in monotonic units, it is not surprising that a sample population with a high degree of nonmonotonicity would display a large degree of spatial selectivity. The results of the present study corroborate the relationship between monotonicity and spatial selectivity, but we did not find a decreased tolerance of spatial tuning to sound level in monotonic units. There may be fundamental differences in processing between awake and anesthetized animals that are more complicated than can be addressed by measures of level tuning.

Thus, it seems clear that there are clear differences between spatial tuning in auditory cortex in awake and anesthetized animals, and these differences parallel those seen in other auditory domains. We therefore suggest an alternate hypothesis to the panoramic view of spatial coding (Middlebrooks et al. 1994): one where each location is equally represented by different populations of neurons. It would be interesting to study the causes that give rise to such differences between spatial tuning in awake and anesthetized animals, but we feel that this is not the most important question. Rather, it is more important to further study the nature of spatial

processing in the full field in awake animals to understand how such heterogeneous spatial tuning functions arise, and how this information is used by higher cognitive processes.

Also in this thesis, we developed a novel approach for estimating spatial receptive fields using 24-speaker simultaneous sound delivery. In natural acoustic environments, the auditory system often has to identify a sound source's location among competing sounds from multiple locations. A study from our laboratory suggested that interactions between spatial locations are strong and suppressive, and with equal strength within and outside a neuron's spatial receptive field (Zhou & Wang 2012). The method used here employed a systems identification approach to estimate spatial receptive fields by simultaneously presenting broad band sounds with randomized intensities from all speakers. Most neurons tested with this stimulus paradigm showed low agreement between the spatial receptive field measured in the 1- and 24-speaker stimulus delivery conditions. We had little success in determining which neural properties led to high agreement between the two conditions, but preliminary results suggested that such neurons may be well suited for coding spatial location in spatially diffuse noise.

7.2 Future studies of spatial coding

Given the observation of substantial spatial receptive field heterogeneity in chapter 3, several questions arise. The first is related to the creation of these various SRF types. What are the underlying auditory cue combinations that neurons are sensitive to? In chapter six, we showed that for some neurons, including those with specific front/rear and elevation preferences, such receptive fields can be predicted to some degree using a linear synthesis of the neuron's spectral weighting function and the animal's head-related transfer function (HRTF). A better method would be to employ closed field acoustic stimulation for a more controlled and accurate measurement of binaural spectral weighting functions. This type of experiment has been performed previously in auditory cortex in anesthetized ferrets (Schnupp et al. 2001; Mrsic-flogel

et al. 2005) and in the nucleus of the brachium of the inferior colliculus in awake marmosets (Slee & Young 2013), but receptive fields in these studies were highly homogenous and not very strongly dependent on detailed spectral cues. It would be extremely interesting to see how spectral cues contribute to spatial receptive fields in auditory cortex in marmosets.

The second question is where does spatial information go from auditory cortex? The distribution of tuning areas we observed suggests that a very large fraction of neurons are spatially tuned, but it is not known what fraction of these neuronal responses are actually decoded by higher auditory areas involved in spatial processing. Given the rapid expansion of neurophysiological tools to investigate functional connections in the brain (e.g. Znamenskiy & Zador 2013), it seems that it may now be possible to more directly study causality of the activity of specific neural populations in spatial perception and behavior.

Finally, the finding that some neurons' receptive fields can be estimated using random spatial profile stimuli warrants further study. As mentioned previously, it is possible that these neurons are involved in spatial coding in noisy environments. Another possible application for RSP stimuli is the study of the coding of acoustic field width. Neurons in auditory cortex have been previously shown using RSS stimuli to have varying preferences for spectral contrast (Barbour & Wang 2003b). It may be possible that neurons have similar preferences for the spatial extent of sound sources.

7.3 Auditory behavior and sound localization in marmosets.

We next described an auditory operant behavior paradigm that is well suited to the study of acoustic perception in the marmoset monkey in which animals can be trained quickly. This paradigm takes advantage of the marmoset's natural licking behavior. Thus far, it has been used to test absolute hearing thresholds (Osmanski & Wang 2011), processing of pitch (Osmanski et al. 2013), and now perceptual acuity for spatial location. We found that marmosets can

discriminate sound location to a degree comparable to other mammals of similar size. This result suggests that marmosets are not innately expert localizers, despite the fact that they occupy a natural habitat in which heightened localization abilities may be beneficial. It was interesting that although marmosets quickly learned the location discrimination task conceptually, subjects seemed to require a significant amount of training to reach a final perceptual threshold. A few subjects even reached a very high, but seemingly stable threshold prior to slowly improving to the final reported threshold. Very long training times have been reported previously, for instance on an ITD discrimination task (Scott et al. 2007) and also on a delayed match to sample task (Fritz, Mishkin, et al. 2005). However, given the fundamental importance of spatial localization in natural environments, the current observation is puzzling. It is possible that marmoset HRTFs are not optimized for localizing in the median plane, as spectral cues in this region are less regular than at lateral positions (Slee & Young 2010). However, it is also possible that the laboratory housing environment is not optimal for the development of such abilities.

A promising feature of the behavior described here is its suitability for pairing with electrophysiological recording. The behavior measurement apparatus and reward delivery system were both designed to work in concert with current single unit recording procedures employed in the lab, and testing has shown that the setup is well suited for this endeavor. This creates the potential for achieving a more complete understanding of acoustic signal processing in the primate brain. Some of the most obvious applications for this task are the perception of vocal acoustics (Miller & Wang 2006; Eliades & Wang 2008) and pitch processing (Bendor & Wang 2005). In chapter 5, we used this simple task to study representation of space while marmosets discriminated spatial location. Successful implementation of an auditory operant conditioning task adds to the existing attractiveness of the marmoset as a model for auditory processing and opens the door to new exciting discoveries in the field of auditory neuroscience.

7.4 Future localization behavior studies in marmosets

The most obvious experiment to conduct at this point is a test of the absolute localization abilities of marmosets. In this type of task, a subject is required to orient to or approach a sound location rather than simply report the change of a sound location. Several mammalian species have been tested in these types of tasks for azimuthal localization, including primates (Waser 1977; Whittington 1981), cats (Casseday & Neff 1973), and ferrets (Parsons et al. 1999). Cats have also been tested in such a task which required orientation to targets which varied in elevation as well as azimuth (May & Huang 1996). One of the benefits of such a task is that, when combined with an acuity task, the specific HRTF cues which enable sound localization can be dissociated from those which allow discrimination of sound location. For instance, in cats, absolute location relies on mid frequency cues (below 18kHz), the region where the "first notch" is located in the HRTF (Rice et al. 1992; Huang & May 1996a). The first notch in marmosets is located between 12 and 24 kHz (Slee & Young 2010).

7.5 The effects of behavioral engagement on spatial responses in auditory cortex

We recorded single-unit responses in auditory cortex of marmosets while they performed a spatial discrimination task in different regions of the full spatial field. Comparing these responses to those measured while marmosets sat passively, we found a subset of neurons with increased firing rates to one or more target locations during task engagement. Effects at background locations were mixed. Firing increases served to increase firing rates relative to the reference location. The observation that firing rate increases occurred at some and not all locations while also distributed throughout receptive fields indicates that effects in this task were largely locations specific. While firing rate increases could occur throughout receptive fields, the tendency for firing rate changes to occur more consistently and strongly at target locations rather than at both target and background locations suggests that these effects acted to tailor spatial

representation for the specific behavior task. We did not find evidence of single trial modulation based in behavioral outcome: effects were not strongly correlated with hit rate, and firing rates to background locations in hit and miss trials were not appreciably different; however we did find a significant difference in spontaneous firing rates in false alarms and correct rejections.

Comparing the behavior effects we observed with those in previous studies, the varied but significant effects of behavioral context on spatial receptive fields in auditory cortex suggest that spatial tuning in the passive condition provides only part of the picture of spatial representation in auditory cortex. As mentioned earlier, these and other behavior studies may reconcile disparate observations regarding spatial tuning in auditory cortex. The observation that large firing rate increases can occur far from best location may be evidence of broader inputs that are masked in the passive awake condition when compared to the anesthetized state. Conversely, another behavior task may lead to still more selective tuning compared to the passive state (Lee & Middlebrooks 2011).

7.6 Future studies of spatial coding in behaving marmosets.

The true nature of spatial representation may not be well understood by studying non-behaving subjects, and a complete picture will likely require further behavior studies. It has become increasingly apparent through time that the effects of task engagement can be highly varied and is often dependent on the exact task structure (Benson & Hienz 1978; Benson et al. 1981; Fritz et al. 2003; Fritz et al. 2007; Atiani et al. 2009; David et al. 2012). It is therefore tempting to suggest a diverse set of spatial tasks in which to measure all the potential ways that auditory cortex may adapt to the various task demands. However, it is possible that more could be learned by choosing the single most ethologically relevant task and fully characterizing the neural mechanisms involved in producing accurate behavior.

A candidate for such a task is three dimensional phonotaxis. Horizontal phonotaxis has been extensively employed in ferrets (Parsons et al. 1999; King & Parsons 1999; Leach et al. 2013) and cats (Malhotra & Lomber 2007; Lomber et al. 2007; Lomber & Malhotra 2008) to study the effects of different developmental and physiological manipulations on sound localization behavior. Marmosets, as an arboreal species, are an ideal model in which to study phonotaxis in three dimensions. Our lab has recently developed a wireless recording system (Roy & Wang 2012) which is currently in use in studies of vocal communication. This system, combined with a spatially and temporally accurate head position tracking system, would enable study of spatial coding in auditory cortex of a marmoset performing a spatial localization task in three dimensions. Ideally, experiments would also not be restricted to the auditory cortex, but extend to the many brain regions which are involved in spatial perception and behavior, such as parietal cortex, which is known to receive projections from the caudal belt areas (Romanski et al. 1999). Only by studying responses in multiple brain areas can one begin to truly understand the computations involved in producing sound localization behavior.

7.7 Differences in spatial processing between areas A1, CM/CL, and R/RT

Several studies have shown that the distributions of spatial tuning properties vary quantitatively, but not qualitatively, between auditory areas along the rostral-caudal axis, with neurons in caudal areas, statistically, displaying higher selectivity for spatial locations than those in rostral areas and primary auditory cortex (Rauschecker & Tian 2000; Stecker et al. 2003; Woods et al. 2006; Lee & Middlebrooks 2013). This report represents the first time that (a) responses to the full field have been recorded in multiple areas in an awake primate, and (b) behavior effects on spatial responses in primates have been directly compared between multiple auditory areas including significant recordings from the caudal belt. The present results mirror previous studies in that the differences in both passive tuning properties and the effects of behavior between these areas are quantitative rather than qualitative. Although it has been previously suggested that increased

levels of nonmonotonic level tuning may account for the differences between caudal and primary auditory cortex, we did not find this to be the case. We did find evidence for a qualitative difference in the effects of behavior, specifically multiplicative in addition to additive gain during discrimination task engagement in areas CM/CL, but this effect was small. Future studies to better understand the different contributions of rostral and caudal areas to sound localization would be similar to those suggested for spatial coding in general: that is testing specifically how spatial receptive fields are constructed in such areas, and their direct influence in spatial perception.

References

Agus, T.R., Thorpe, S.J. & Pressnitzer, D., 2010. Rapid formation of robust auditory memories: insights from noise. *Neuron*, 66(4), pp.610–18.

Aitkin, L.M. et al., 1986. Frequency representation in auditory cortex of the common marmoset (Callithrix jacchus jacchus). *The Journal of comparative neurology*, 252(2), pp.175–85.

Aitkin, L.M. & Jones, R., 1992. Azimuthal processing in the posterior auditory thalamus of cats. *Neuroscience letters*, 142(1), pp.81–4.

Atiani, S. et al., 2009. Task difficulty and performance induce diverse adaptive patterns in gain and shape of primary auditory cortical receptive fields. *Neuron*, 61(3), pp.467–80.

Barbour, D.L., 2002. *Coding of complex temporal and spectral features in the auditory cortex of awake primates*. Johns Hopkins University School of Medicine.

Barbour, D.L. & Wang, X., 2003a. Auditory cortical responses elicited in awake primates by random spectrum stimuli. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 23(18), pp.7194–206.

Barbour, D.L. & Wang, X., 2003b. Contrast tuning in auditory cortex. *Science*, 299(5609), pp.1073–5.

Barone, P. et al., 1996. Cortical synthesis of azimuth-sensitive single-unit responses with nonmonotonic level tuning: a thalamocortical comparison in the cat. *Journal of Neurophysiology*, 75, pp.1206–1220.

Bartlett, E. & Wang, X., 2005. Long-lasting modulation by stimulus context in primate auditory cortex. *Journal of neurophysiology*, 94(1), pp.83–104.

Bartlett, E.L. & Wang, X., 2007. Neural representations of temporally modulated signals in the auditory thalamus of awake primates. *Journal of neurophysiology*, 97(2), pp.1005–17.

Bendor, D. & Wang, X., 2007. Differential neural coding of acoustic flutter within primate auditory cortex. *Nature neuroscience*, 10(6), pp.763–71.

Bendor, D. & Wang, X., 2008. Neural response properties of primary, rostral, and rostrotemporal core fields in the auditory cortex of marmoset monkeys. *Journal of neurophysiology*, 100(2), pp.888–906.

Bendor, D. & Wang, X., 2005. The neuronal representation of pitch in primate auditory cortex. *Nature*, 436(7054), pp.1161–5.

Benson, D.A. & Hienz, R.D., 1978. Single-unit activity in the auditory cortex of monkeys selectively attending left vs. right ear stimuli. *Brain research*, 159(2), pp.307–320.

Benson, D.A., Hienz, R.D. & Goldstein, M.H., 1981. Single-unit activity in the auditory cortex of monkeys actively localizing sound sources: spatial tuning and behavioral dependency. *Brain research*, 219(2), pp.249–67.

Bizley, J.K. et al., 2007. Role of auditory cortex in sound localization in the midsagittal plane. *Journal of neurophysiology*, 98(3), pp.1763–74.

Blauert, J., 1997. *Spatial hearing: the psychophysics of human sound localization* 3rd ed., Cambridge, MA: MIT Press.

Bouchardet da Fonesca, G.A. & Lacher, T.E., 1984. Exudate-feeding by Callithrixjacchus penicillata in. *Primates*, 25(4), pp.441–450.

Bremen, P., van Wanrooij, M.M. & van Opstal, a J., 2010. Pinna cues determine orienting response modes to synchronous sounds in elevation. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 30(1), pp.194–204.

Britten, K. & Newsome, W., 1996. A relationship between behavioral choice and the visual responses of neurons in macaque MT. *Visual Neuroscience*, 13, pp.87–100.

Brosch, M., Selezneva, E. & Scheich, H., 2005. Nonauditory events of a behavioral procedure activate auditory cortex of highly trained monkeys. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 25(29), pp.6797–806.

Brown, C.H. et al., 1980. Localization of noise bands by Old World monkeys. *The Journal of the Acoustical Society of America*, 68(1), pp.127–132.

Brown, C.H. et al., 1982. Vertical and horizontal sound localization in primates. *Journal of the Acoustical Socieity of America*, 72(6), pp.1804–1811.

Brown, C.H. & May, B.J., 2005. Comparative mammalian sound localization. In A. N. Popper & R. R. Fay, eds. *Sound Source Localization*. Springer, pp. 124–178.

Brugge, J., Reale, R. & Hind, J., 1994. Simulation of free-field sound sources and its application to studies of cortical mechanisms of sound localization in the cat. *Hearing research*, 73(1), pp.67–84.

Brugge, J.F., Reale, R.A. & Hind, J.E., 1996. The Structure of Spatial Receptive Fields of Neurons in Primary Auditory Cortex of the Cat. *The Journal of Neuroscience*, 16(14), pp.4420 – 4437.

Butler, R., 1986. The bandwidth effect on monaural and binaural localization. *Hearing research*, 21(1), pp.67–73.

Cai, R. et al., 2009. Environmental enrichment improves behavioral performance and auditory spatial representation of primary auditory cortical neurons in rat. *Neurobiology of learning and memory*, 91(4), pp.366–76.

Campbell, R. a a et al., 2006. Binaural-level functions in ferret auditory cortex: evidence for a continuous distribution of response properties. *Journal of neurophysiology*, 95(6), pp.3742–55.

Casseday, J. & Neff, W., 1973. Localization of pure tones. *The Journal of the Acoustical Society of America*, 54(2), pp.365–372.

Cohen, M. a, Horowitz, T.S. & Wolfe, J.M., 2009. Auditory recognition memory is inferior to visual recognition memory. *Proceedings of the National Academy of Sciences of the United States of America*, 106(14), pp.6008–10.

Colombo, M. & D'Amato, M., 1986. A comparison of visual and auditory short-term memory in monkeys (Cebus apella). *The Quarterly journal of experimental psychology section B: comparative and physiological psychology*, 38(4), pp.425–448.

D'Amato, M., 1988. A search for tonal pattern perception in Cebus monkeys: Why monkeys can't hum a tune. *Music Perception*, 5(4), pp.453–480.

David, S. V, Fritz, J.B. & Shamma, S. a, 2012. Task reward structure shapes rapid receptive field plasticity in auditory cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 109(6), pp.2144–9.

Delgutte, B. & Joris, P., 1999. Receptive fields and binaural interactions for virtual-space stimuli in the cat inferior colliculus. *Journal of neurophysiology*, 81(6), pp.2833–2851.

Dodd, J. V et al., 2001. Perceptually bistable three-dimensional figures evoke high choice probabilities in cortical area MT. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 21(13), pp.4809–21.

Dong, C. et al., 2013. Behavioral Modulation of Neural Encoding of Click-Trains in the Primary and Nonprimary Auditory Cortex of Cats. *Journal of Neuroscience*, 33(32), pp.13126–13137.

Efrati, A. & Gutfreund, Y., 2011. Early life exposure to noise alters the representation of auditory localization cues in the auditory space map of the barn owl. *Journal of neurophysiology*, 105(5), pp.2522–35.

Eisenman, L.M., 1974. Neural encoding of sound location: An electrophysiological study in auditory cortex (A1) of the cat using free field stimuli. *Brain research*, 75(2), pp.203–214.

Eliades, S.J. & Wang, X., 2008. Neural substrates of vocalization feedback monitoring in primate auditory cortex. *Nature*, 453(7198), pp.1102–6.

Frey, H., 2003. Cats can detect repeated noise stimuli. *Neuroscience Letters*, 346(1-2), pp.45–48.

Fritz, J.B. et al., 2003. Rapid task-related plasticity of spectrotemporal receptive fields in primary auditory cortex. *Nature neuroscience*, 6(11), pp.1216–23.

Fritz, J.B., Elhilali, M. & Shamma, S., 2005. Active listening: task-dependent plasticity of spectrotemporal receptive fields in primary auditory cortex. *Hearing research*, 206(1-2), pp.159–76.

Fritz, J.B., Elhilali, M. & Shamma, S.A., 2007. Adaptive changes in cortical receptive fields induced by attention to complex sounds. *Journal of neurophysiology*, 98(4), pp.2337–46.

Fritz, J.B., Mishkin, M. & Saunders, R.C., 2005. In search of an auditory engram. *Proceedings of the National Academy of Sciences of the United States of America*, 102(26), pp.9359–64.

Gescheider, G.A., 1985. Psychophysical Measurement of Thresholds. In *Psychophysics: Method, Theory, and Application*. pp. 1–37.

Goldberg, J.M. & Brown, P.B., 1969. Response of binaural neurons of dog superior olivary complex to dichotic tonal stimuli: some physiological mechanisms of sound localization. *Journal of Neurophysiology*, 32(4), pp.613–36.

Goossens, T., van de Par, S. & Kohlrausch, A., 2008. On the ability to discriminate Gaussian-noise tokens or random tone-burst complexes. *The Journal of the Acoustical Society of America*, 124(4), pp.2251–62.

Greene, T.C., 1929. The ability to localize sound: a study of binaural hearing in patients with tumor of the brain. *Arch Surg*, 18(4), pp.1825–1841.

Guttman, N. & Julesz, B., 1963. Lower limits of auditory periodicity analysis. *The Journal of the Acoustical Society of America*, 35(4), p.610.

Hackett, T.A., 2011. Information flow in the auditory cortical network. *Hearing Research*, 271(1), pp.133–146.

Hanna, T.E., 1984. Discrimination of reproducible noise as a function of bandwidth and duration. *Perception & psychophysics*, 36(5), pp.409–16.

Harnett, M.T. et al., 2013. Potassium Channels Control the Interaction between Active Dendritic Integration Compartments in Layer 5 Cortical Pyramidal Neurons. *Neuron*, 79(3), pp.516–529.

Harrison, J.M. & Downey, P., 1970. Intensity changes at the ear as a function of the azimuth of a tone source: a comparative study. *The Journal of the Acoustical Society of America*, 47(6), pp.1509–18.

Hasselmo, M. & Schnell, E., 1994. Laminar selectivity of the cholinergic suppression of synaptic transmission in rat hippocampal region CA1: computational modeling and brain slice physiology. *The Journal of neuroscience*, 14(6), pp.3898–3914.

Heffner, H. & Masterton, R.B., 1980. Hearing in Glires: Domestic rabbit, cotton rat, feral house mouse, and kangaroo rat. *The Journal of the Acoustical Society of America*, 68(6), p.1584.

Heffner, H.E., 1997. The role of macaque auditory cortex in sound localization. *Acta oto-laryngologica. Supplementum*, 117(s532), pp.22–7.

Heffner, H.E. & Heffner, R.R.S., 1985. Sound localization in wild Norway rats (Rattus norvegicus). *Hearing Research*, 19(2), pp.151–155.

Heffner, H.E. & Heffner, R.S., 1990. Effect of bilateral auditory cortex lesions on sound localization in Japanese macaques. *Journal of neurophysiology*, 64(3), pp.915–31.

Heffner, H.E.H.E. & Heffner, R.S.R.S., 1984. Sound Localization in Large Mammals: Localization of Complex Sounds by Horses. *Behavioral neuroscience*, 98(3), pp.541–555.

Heffner, Rickye S & Heffner, H.E., 1992. Hearing in Large Mammals: Sound-Localization Acuity in Cattle (Bos taurus) and Goats(Capra hircus). *Journal of comparative psychology*, 106(2), pp.107–113.

Heffner, R.S. & Heffner, H.E., 1988a. Sound localization acuity in the cat: effect of azimuth, signal duration, and test procedure. *Hearing research*, 36(2-3), pp.221–32.

Heffner, R.S. & Heffner, H.E., 1988b. Sound localization and use of binaural cues by the gerbil (Meriones unguiculatus). *Behavioral neuroscience*, 102(3), pp.422–8.

Heffner, R S & Heffner, H.E., 1992. Visual factors in sound localization in mammals. *The Journal of comparative neurology*, 317(3), pp.219–32.

Heffner, R.S., Heffner, H.E. & Koay, G., 1995. Sound localization in chinchillas. II. Front/back and vertical localization. *Hearing research*, 88(1-2), pp.190–8.

Heffner, R.S.R.S. & Heffner, H.E.H.E., 1989. Sound localization, use of binaural cues and the superior olivary complex in pigs. *Brain, behavior and evolution*, 33(4), pp.248–58.

Heil, P., Rajan, R. & Irvine, D.R., 1992. Sensitivity of neurons in cat primary auditory cortex to tones and frequency-modulated stimuli. I: Effects of variation of stimulus parameters. *Hearing research*, 63(1-2), pp.108–34.

Houweling, A.R. & Brecht, M., 2008. Behavioural report of single neuron stimulation in somatosensory cortex. *Nature*, 451(7174), pp.65–8.

Huang, a Y. & May, B.J., 1996a. Sound orientation behavior in cats. II. Mid-frequency spectral cues for sound localization. *The Journal of the Acoustical Society of America*, 100(2), pp.1070–80.

Huang, a Y. & May, B.J., 1996b. Spectral cues for sound localization in cats: effects of frequency domain on minimum audible angles in the median and horizontal planes. *The Journal of the Acoustical Society of America*, 100(4), pp.2341–8.

Imig, T.J. et al., 1990. Single-unit selectivity to azimuthal direction and sound pressure level of noise bursts in cat high-frequency primary auditory cortex. *Journal of Neurophysiology*, 63(6), pp.1448–1466.

Jazayeri, M., Lindbloom-Brown, Z. & Horwitz, G.D., 2012. Saccadic eye movements evoked by optogenetic activation of primate V1. *Nature neuroscience*, 15(10), pp.1368–70.

Jazayeri, M. & Movshon, J.A., 2006. Optimal representation of sensory information by neural populations. *Nature neuroscience*, 9(5), pp.690–6.

Jenkins, W.M. & Masterton, R.B., 1982. Sound localization: effects of unilateral lesions in central auditory system. *Journal of neurophysiology*, 47(6), pp.987–1016.

Jenkins, W.M. & Merzenich, M.M., 1984. Role of cat primary auditory cortex for sound-localization behavior. *Journal of neurophysiology*, 52(5), pp.819–47.

Johnson, D., 1980. The relationship between spike rate and synchrony in responses of auditory-nerve fibers to single tones. *The Journal of the Acoustical Society of America*, 68(4), pp.1115–1122.

Kaas, J.H. & Hackett, T. a, 1998. Subdivisions of auditory cortex and levels of processing in primates. *Audiology & neuro-otology*, 3(2-3), pp.73–85.

Kaas, J.H. & Hackett, T. a, 2000. Subdivisions of auditory cortex and processing streams in primates. *Proceedings of the National Academy of Sciences of the United States of America*, 97(22), pp.11793–9.

Kadia, S.C. & Wang, X., 2003. Spectral integration in A1 of awake primates: neurons with single- and multipeaked tuning characteristics. *Journal of neurophysiology*, 89(3), pp.1603–22.

Kaernbach, C., 1993. Temporal and spectral basis of the features perceived in repeated noise. *Journal of the Acoustical Socieity of America*, 94(1), pp.91–97.

Kaernbach, C., 2004. The Memory of Noise. *Experimental Psychology (formerly "Zeitschrift für Experimentelle Psychologie")*, 51(4), pp.240–248.

Kaernbach, C. & Schulze, H., 2002. Auditory sensory memory for random waveforms in the Mongolian gerbil. *Neuroscience letters*, 329(1), pp.37–40.

Kajikawa, Y. et al., 2008. Coding of FM sweep trains and twitter calls in area CM of marmoset auditory cortex. *Hearing research*, 239(1-2), pp.107–25.

Kaminski, J., Call, J. & Fischer, J., 2004. Word learning in a domestic dog: evidence for "fast mapping". *Science*, 304(5677), pp.1682–3.

Kelly, J.B. & Kavanagh, G.L., 1987. Localization by the Ferret (Mustela putorius). *Journal of Neurophysiology*, 57(6).

Kilgard, M.P. & Merzenich, M.M., 1999. Distributed representation of spectral and temporal information in rat primary auditory cortex. *Hearing research*, 134(1-2), pp.16–28.

King, a J. & Parsons, C.H., 1999. Improved auditory spatial acuity in visually deprived ferrets. *The European journal of neuroscience*, 11(11), pp.3945–56.

Klein, D.J. et al., 1997. Robust spectrotemporal reverse correlation for the auditory system: optimizing stimulus design. *Journal of computational neuroscience*, 9(1), pp.85–111.

Klumpp, R.G. & Eady, H.R., 1956. Some measurements of interaural time difference thresholds. *Journal of the Acoustical Society of America*, 28, pp.859–860.

Kretzschmar, C. et al., 2008. Echoic memory in pigeons. *Behavioural processes*, 79(2), pp.105–10.

De la Mothe, L. a et al., 2012. Cortical connections of auditory cortex in marmoset monkeys: lateral belt and parabelt regions. *Anatomical record (Hoboken, N.J. : 2007)*, 295(5), pp.800–21.

De la Mothe, L. a et al., 2006a. Cortical connections of the auditory cortex in marmoset monkeys: core and medial belt regions. *The Journal of comparative neurology*, 496(1), pp.27–71.

De la Mothe, L. a et al., 2006b. Thalamic connections of the auditory cortex in marmoset monkeys: core and medial belt regions. *The Journal of comparative neurology*, 496(1), pp.72–96.

Leach, N.D. et al., 2013. Cortical Cholinergic Input Is Required for Normal Auditory Perception and Experience-Dependent Plasticity in Adult Ferrets. *Journal of Neuroscience*, 33(15), pp.6659–6671.

Leakey, D.M., 1959. Some Measurements on the Effects of Interchannel Intensity and Time Differences in Two Channel Sound Systems. *The Journal of the Acoustical Society of America*, 31(7), pp.977–986.

Lee, C.-C. & Middlebrooks, J.C., 2011. Auditory cortex spatial sensitivity sharpens during task performance. *Nature Neuroscience*, 14(1), pp.108–114.

Lee, C.-C. & Middlebrooks, J.C., 2013. Specialization for sound localization in fields A1, DZ, and PAF of cat auditory cortex. *Journal of the Association for Research in Otolaryngology : JARO*, 14(1), pp.61–82.

Lomber, S.G. & Malhotra, S., 2008. Double dissociation of "what" and "where" processing in auditory cortex. *Nature neuroscience*, 11(5), pp.609–16.

Lomber, S.G., Malhotra, S. & Hall, A.J., 2007. Functional specialization in non-primary auditory cortex of the cat: areal and laminar contributions to sound localization. *Hearing research*, 229(1-2), pp.31–45.

Lu, T., Liang, L. & Wang, X, 2001. Neural representations of temporally asymmetric stimuli in the auditory cortex of awake primates. *Journal of neurophysiology*, 85(6), pp.2364–80.

Lu, T., Liang, L. & Wang, Xiaoqin, 2001. Temporal and rate representations of time-varying signals in the auditory cortex of awake primates. *Nature neuroscience*, 4(11), pp.1131–8.

Macpherson, E. a & Middlebrooks, J.C., 2000. Localization of brief sounds: effects of level and background noise. *The Journal of the Acoustical Society of America*, 108(4), pp.1834–49.

Malhotra, S. & Lomber, S.G., 2007. Sound localization during homotopic and heterotopic bilateral cooling deactivation of primary and nonprimary auditory cortical areas in the cat. *Journal of neurophysiology*, 97(1), pp.26–43.

Martin, R.L. & Webster, W.R., 1987. The auditory spatial acuity of the domestic cat in the interaural and median vertical planes horizontal. *Hearing Research*, 30(2-3), pp.239–252.

Masterton, B. et al., 1975. Neuroanatomical basis of binaural phase-difference analysis for sound localization: a comparative study. *Journal of comparative and physiological psychology*, 89(5), pp.379–386.

May, B.J. & Huang, a Y., 1996. Sound orientation behavior in cats. I. Localization of broadband noise. *The Journal of the Acoustical Society of America*, 100(2), pp.1059–69.

Melcher, D. & Colby, C.L., 2008. Trans-saccadic perception. *Trends in cognitive sciences*, 12(12), pp.466–73.

Merzenich, M. & Brugge, J.F., 1973. Representation of the cochlear partition on the superior temporal plane of the macaque monkey. *Brain research*, 50(2), pp.275–296.

Mickey, B.J. & Middlebrooks, J.C., 2003. Representation of auditory space by cortical neurons in awake cats. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 23(25), pp.8649–63.

Middlebrooks, J.C. et al., 1994. A panoramic code for sound location by cortical neurons. *Science*, 264(5160), pp.842–4.

Middlebrooks, J.C. et al., 2002. Location Signaling by Cortical Neurons. In D. Oertel, R. R. Fay, & A. N. Popper, eds. *Integrative Functions in the Mammaliann Auditory Pathway*. New York, NY: Springer, pp. 319–357.

Middlebrooks, J.C. & Bremen, P., 2013. Spatial Stream Segregation by Auditory Cortical Neurons. *Journal of Neuroscience*, 33(27), pp.10986–11001.

Middlebrooks, J.C. & Pettigrew, J.D., 1981. Functional classes of neurons in primary auditory cortex of the cat distinguished by sensitivity to sound location. *Journal of Neuroscience*, 1(1), pp.107–120.

Miles, R.C. & Meyer, D.R., 1956. Learning sets in marmosets. *Journal of comparative and physiological psychology*, 49(3), pp.219–222.

Miller, C.T. et al., 2009. Antiphonal call timing in marmosets is behaviorally significant: interactive playback experiments. *Journal of comparative physiology. A, Neuroethology, sensory, neural, and behavioral physiology*, 195(8), pp.783–9.

Miller, C.T., Dimauro, A., et al., 2010. Vocalization Induced CFos Expression in Marmoset Cortex. *Frontiers in integrative neuroscience*, 4, pp.1–15.

Miller, C.T., Mandel, K. & Wang, X., 2010. The communicative content of the common marmoset phee call during antiphonal calling. *American journal of primatology*, 72(11), pp.974–80.

Miller, C.T. & Wang, X., 2006. Sensory-motor interactions modulate a primate vocal behavior: antiphonal calling in common marmosets. *Journal of comparative physiology. A, Neuroethology, sensory, neural, and behavioral physiology*, 192(1), pp.27–38.

Miller, J.M. et al., 1972. Single cell activity in the auditory cortex of Rhesus monkeys: behavioral dependency. *Science*, 177(4047), pp.449–451.

Mrsic-flogel, T.D. et al., 2005. Encoding of Virtual Acoustic Space Stimuli by Neurons in Ferret Primary Auditory Cortex. *Journal of neurophysiology*, 93(6), pp.3489 –3503.

Nakamura, T. et al., 1999. Synergistic release of Ca2+ from IP3-sensitive stores evoked by synaptic activation of mGluRs paired with backpropagating action potentials. *Neuron*, 24(3), pp.727–37.

Nelson, P.C., Smith, Z.M. & Young, E.D., 2009. Wide-dynamic-range forward suppression in marmoset inferior colliculus neurons is generated centrally and accounts for perceptual masking. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 29(8), pp.2553–62.

Nienborg, H. & Cumming, B.G., 2006. Macaque V2 neurons, but not V1 neurons, show choice-related activity. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 26(37), pp.9567–78.

Niwa, M. et al., 2012. Active engagement improves primary auditory cortical neurons' ability to discriminate temporal modulation. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 32(27), pp.9323–34.

Oldfield, S.R. & Parker, S.P., 1984. Acuity of sound localisation: a topography of auditory space. I. Normal hearing conditions. *Perception*, 13(5), pp.581–600.

Osmanski, M., Song, X. & Wang, X., 2013. The Role of Harmonic Resolvability in Pitch Perception in a Vocal Nonhuman Primate, the Common Marmoset (Callithrix jacchus). *The Journal of Neuroscience*, 33(21), pp.9161–9168.

Osmanski, M.S. & Wang, X., 2011. Measurement of absolute auditory thresholds in the common marmoset (Callithrix jacchus). *Hearing research*, 277(1-2), pp.127–33.

Otazu, G.H. et al., 2009. Engaging in an auditory task suppresses responses in auditory cortex. *Nature neuroscience*, 12(5), pp.646–54.

Pan, Y. et al., 2011. Developmentally degraded directional selectivity of the auditory cortex can be restored by auditory discrimination training in adults. *Behavioural brain research*, 225(2), pp.596–602.

Parsons, C.H. et al., 1999. Effects of altering spectral cues in infancy on horizontal and vertical sound localization by adult ferrets. *Journal of neurophysiology*, 82(5), pp.2294–309.

Peña, J.L. & Konishi, M., 2001. Auditory spatial receptive fields created by multiplication. *Science (New York, N.Y.)*, 292(5515), pp.249–52.

Pfingst, B. & O'Connor, T., 1981. Characteristics of neurons in auditory cortex of monkeys performing a simple auditory task. *Journal of neurophysiology*, 45(1), pp.16–34.

Phillips, D. & Orman, S., 1984. Responses of single neurons in posterior field of cat auditory cortex to tonal stimulation. *Journal of neurophysiology*, 51(1), pp.147–163.

Rauschecker, J., 1998. Parallel Processing in the Auditory Cortex of Primates. *Audiology & Neuro-Otology*, 3(2-3), pp.86–103.

Rauschecker, J.P. & Tian, B., 2000. Mechanisms and streams for processing of "what" and "where" in auditory cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 97(22), pp.11800–6.

Ravizza, R., Ravizza, J. & Masterton, R.B., 1972. Contribution of Neocortex to Sound Localization in Opossum (Didelphis virginiana). *Journal of Neurophysiology*, 35(3), pp.344–56.

Recanzone, G.H. et al., 2000. Correlation between the activity of single auditory cortical neurons and sound-localization behavior in the macaque monkey. *Journal of neurophysiology*, 83(5), pp.2723–39.

Remington, E.D., Osmanski, M.S. & Wang, X., 2012. An Operant Conditioning Method for Studying Auditory Behaviors in Marmoset Monkeys A. Claude, ed. *PLoS ONE*, 7(10), p.e47895.

Rendall, D., Rodman, P.S. & Emond, R.E., 1996. Vocal recognition of individuals and kin in free-ranging rhesus monkeys. *Animal Behaviour*, 51(5), pp.1007–1015.

Reser, D.H. et al., 2009. Connections of the marmoset rostrotemporal auditory area: express pathways for analysis of affective content in hearing. *The European journal of neuroscience*, 30(4), pp.578–92.

Rice, J. et al., 1992. Pinna-based spectral cues for sound localization in cat. *Hearing research*, 58(2), pp.132–152.

Roberts, A.C., Robbins, T.W. & Everitt, B.J., 1988. The effects of intradimensional and extradimensional shifts on visual discrimination learning in humans and non-human primates. *The Quarterly journal of experimental psychology section B: comparative and physiological psychology*, 40B(4), pp.321–341.

Romanski, L.M. et al., 1999. Dual streams of auditory afferents target multiple domains in the primate prefrontal cortex. *Nature neuroscience*, 2(12), pp.1131–6.

Rose, J. & Brugge, J., 1967. Phase-locked response to low-frequency tones in single auditory nerve fibers of the squirrel monkey. *Journal of ...*, 30(4), pp.769–793.

Roy, S. & Wang, X., 2012. Wireless multi-channel single unit recording in freely moving and vocalizing primates. *Journal of neuroscience methods*, 203(1), pp.28–40.

Sadagopan, S. & Wang, X., 2008. Level invariant representation of sounds by populations of neurons in primary auditory cortex. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 28(13), pp.3415–26.

Sadagopan, S. & Wang, X., 2009. Nonlinear spectrotemporal interactions underlying selectivity for complex sounds in auditory cortex. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 29(36), pp.11192–202.

Sasaki, E. et al., 2009. Generation of transgenic non-human primates with germline transmission. *Nature*, 459(7246), pp.523–7.

Schnupp, J.W., Mrsic-Flogel, T.D. & King, a J., 2001. Linear processing of spatial cues in primary auditory cortex. *Nature*, 414(6860), pp.200–4.

Schreiner, C., Read, H. & Sutter, M., 2000. Modular organization of frequency integration in primary auditory cortex. *Annual review of neuroscience*, 23, pp.501–529.

Scott, B.H., Malone, B.J. & Semple, M.N., 2007. Effect of behavioral context on representation of a spatial cue in core auditory cortex of awake macaques. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 27(24), pp.6489–99.

Seiden, H.R., 1957. *Auditory acuity of the marmoset monkey (Hapale jacchus)*. Princeton University.

Seyfarth, R., Cheney, D. & Marler, P., 1980. Monkey responses to three different alarm calls: evidence of predator classification and semantic communication. *Science*, 210(4471), pp.801–803.

Sinnott, J.M., Petersen, M.R. & Hopp, S.L., 1985. Frequency and intensity discrimination in humans and monkeys. *The Journal of the Acoustical Society of America*, 78(6), pp.1977–85.

Slee, S.J. & Young, E.D., 2013. Linear Processing of Interaural Level Difference Underlies Spatial Tuning in the Nucleus of the Brachium of the Inferior Colliculus. *Journal of Neuroscience*, 33(9), pp.3891–3904.

Slee, S.J. & Young, E.D., 2010. Sound localization cues in the marmoset monkey. *Hearing research*, 260(1-2), pp.96–108.

Snow, W.B., 1954. Effect of Arrival Time on Stereophonic Localization. *The Journal of the Acoustical Society of America*, 26(6), pp.1071–1074.

Sovijärvi, A. & Hyvärinen, J., 1974. Auditory cortical neurons in the cat sensitive to the direction of sound source movement. *Brain research*, 73(3), pp.455–471.

Spitzer, H., Desimone, R. & Moran, J., 1988. Increased attention enhances both behavioral and neuronal performance. *Science*, 240(4850), pp.338–340.

Spitzer, H. & Richmond, B., 1991. Task difficulty: ignoring, attending to, and discriminating a visual stimulus yield progressively more activity in inferior temporal neurons. *Experimental Brain Research*, 83(2), pp.340–348.

Stabler, S., 1996. Temporal and mean rate discharge patterns of single units in the dorsal cochlear nucleus of the anesthetized guinea pig. *Journal of neurophysiology*, 76(3), pp.1667–1688.

Stecker, G.C. et al., 2003. Spatial sensitivity in field PAF of cat auditory cortex. *Journal of neurophysiology*, 89(6), pp.2889–903.

Stecker, G.C., Harrington, I. a & Middlebrooks, J.C., 2005. Location coding by opponent neural populations in the auditory cortex. *PLoS biology*, 3(3), pp.520–528.

Stepien, L., Cordeau, J. & Rasmussen, T., 1960. The effect of temporal lobe and hippocampal lesions on auditory and visual recent memory in monkeys. *Brain: A Journal of Neurology*, 83, pp.470–489.

Stevens, S.S. & Newman, E.B., 1936. The localization of actual sources of sound. *The American Journal of Psychology*, 48(2), pp.297–306.

Suga, N., O'Neill, W. & Manabe, T., 1979. Harmonic-sensitive neurons in the auditory cortex of the mustache bat. *Science*, 203(4377), pp.270–274.

Swets, J.A., 1961. Decision process in perception. *Psychological Review*, 68(5), pp.301–340.

Taaseh, N., Yaron, A. & Nelken, I., 2011. Stimulus-specific adaptation and deviance detection in the rat auditory cortex. *PloS one*, 6(8), p.e23369.

Tian, B. et al., 2001. Functional specialization in rhesus monkey auditory cortex. *Science*, 292(5515), pp.290–3.

Tibshirani, R., Walther, G. & Hastie, T., 2001. Estimating the number of clusters in a data set via the gap statistic. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(2), pp.411–423.

Treue, S. & Martínez Trujillo, J.C., 1999. Feature-based attention influences motion processing gain in macaque visual cortex. *Nature*, 399(6736), pp.575–9.

Ulanovsky, N. et al., 2004. Multiple time scales of adaptation in auditory cortex neurons. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 24(46), pp.10440–53.

Wang, X., 2000. On cortical coding of vocal communication sounds in primates. *Proceedings of the National Academy of Sciences of the United States of America*, 97(22), pp.11843–9.

Van Wanrooij, M.M. & Van Opstal, a J., 2004. Contribution of head shadow and pinna cues to chronic monaural sound localization. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 24(17), pp.4163–71.

Warren, R.M. et al., 2001. Detection of acoustic repetition for very long stochastic patterns. *Perception & psychophysics*, 63(1), pp.175–82.

Waser, P.M., 1977. Sound localization by monkeys: A field experiment. *Behavioral Ecology and Sociobiology*, 2(4), pp.427–431.

Watkins, P. V & Barbour, D.L., 2011. Rate-level responses in awake marmoset auditory cortex. *Hearing research*, 275(1-2), pp.30–42.

Watkins, P. V & Barbour, D.L., 2008. Specialized neuronal adaptation for preserving input sensitivity. *Nature neuroscience*, 11(11), pp.1259–61.

Whittington, D., 1981. Eye and head movements to auditory targets. *Experimental Brain Research*, 41(3-4), pp.358–363.

Wightman, F.L. & Kistler, D.J., 1989. Headphone simulation of free-field listening. I: Stimulus synthesis. *The Journal of the Acoustical Society of America*, 85(2), pp.858–67.

Womelsdorf, T. et al., 2006. Dynamic shifts of visual receptive fields in cortical area MT by spatial attention. *Nature neuroscience*, 9(9), pp.1156–60.

Woods, T.M. et al., 2006. Effects of stimulus azimuth and intensity on the single-neuron activity in the auditory cortex of the alert macaque monkey. *Journal of neurophysiology*, 96(6), pp.3323–37.

Wortis, B. & Pfeffer, A.Z., 1948. Unilateral Auditory-Spatial Agnosia. *Journal of Nervous & Mental Disease*, 108(3), pp.181–186.

Xu, L. et al., 1998. Sensitivity to Sound-Source Elevation in Nontonotopic Auditory Cortex. *Journal of neurophysiology*, 80(2), pp.882–894.

Xu, N. et al., 2012. Nonlinear dendritic integration of sensory and motor input during an active sensing task. *Nature*, 492(7428), pp.247–251.

Yang, Y. et al., 2008. Millisecond-scale differences in neural activity in auditory cortex can drive decisions. *Nature neuroscience*, 11(11), pp.1262–3.

Yao, J.D., Bremen, P. & Middlebrooks, J.C., 2013. Rat Primary Auditory Cortex is Tuned Exclusively to the Contralateral Hemifield. *Journal of neurophysiology*, (949).

Young, E.D. & Calhoun, B.M., 2005. Nonlinear modeling of auditory-nerve rate responses to wideband stimuli. *Journal of neurophysiology*, 94(6), pp.4441–54.

Yu, J., 2003. *Spectral information encoding in the cochlear nucleus and inferior colliculus: a study based on the random spectral shape method*. Johns Hopkins University School of Medicine.

Yu, J.J. & Young, E.D., 2000. Linear and nonlinear pathways of spectral information transmission in the cochlear nucleus. *Proceedings of the National Academy of Sciences of the United States of America*, 97(22), pp.11780–6.

Zatorre, R. & Penhune, V., 2001. Spatial localization after excision of human auditory cortex. *The Journal of Neuroscience*, 21(16), pp.6321–6328.

Zhou, B., Green, D.M. & Middlebrooks, J.C., 1992. Characterization of external ear impulse responses using Golay codes. *The Journal of the Acoustical Society of America*, 92(2), pp.1169–71.

Zhou, Y. & Wang, X., 2012. Level dependence of spatial processing in the primate auditory cortex. *Journal of neurophysiology*, 108(3), pp.810–26.

Znamenskiy, P. & Zador, A.M., 2013. Corticostriatal neurons in auditory cortex drive decisions during auditory discrimination. *Nature*, 497(7450), pp.482–5.

Curriculum Vitae – Evan Remington

11/2/2013

Educational History

Ph.D. (expected 2013) Biomedical Engineering, Johns Hopkins University School of Medicine
B.S. (2006) Biomedical Engineering, Case Western Reserve University

Other Professional Experience

St. Jude Medical Cardiac Rhythm management Division (2005-2006)
Student Internship, software engineering - Cardiac Rhythm Management Division
Software requirements and algorithm design for implantable cardioverter-defribrillators (ICD)

Case Western Reserve University (2003-2004)
Research Assistant, Department of Biomedical Engineering
In vitro models of nonsynaptic epileptiform activity

Honors and awards

Provost's Scholarship (CWRU, 2001-2005)

Publications

Remington, E. D., Osmanski, M. S., & Wang, X. (2012). An Operant Conditioning Method for Studying Auditory Behaviors in Marmoset Monkeys. (A. Claude, Ed.) *PLoS ONE*, *7*(10), e47895. doi:10.1371/journal.pone.0047895

Conferences and Seminars

Remington E and Wang X. Task-dependent spatial responses in auditory cortex of common marmosets. Program 926. 2013 Abstracts of The Thirty-Sixth Annual Midwinter Research Meeting of the Association for Research in Autolaryngology. Baltimore, MD: Association for Research in Autolaryngology, 2013. (Poster)

Remington E and Wang X. Task-dependent spatial responses in auditory cortex of common marmosets. 2012 Neuroscience Meeting Planner. New Orleans, MS: Society for Neuroscience, 2012 (Poster)

Remington E and Wang X. Spatial processing in marmoset auditory cortex. Beijing International Workshop on Auditory Neuroscience. 2012. (Invited talk)

Remington E and Wang X. Spatial hearing acuity of the common marmoset monkey. Program No. 957. 2011 Abstracts of The Thirty-Fourth Annual Midwinter Research Meeting of the Association for Research in Otolaryngology. Baltimore, MD: Association for Research in Autolaryngology, 2011. (Poster)

Remington E and Wang X. Spatial auditory processing in auditory cortex of sleeping marmosets. Program No. 556.4/BB12. 2009 Neuroscience Meeting Planner. Chicago, IL: Society for Neuroscience, 2009. (Poster)